

Carbohydrate Active enZYmes: functional study and applicative perspectives

Roberta Iacono



Carbohydrate Active enZYmes: functional study and applicative perspectives

Università degli Studi di Napoli Federico II
Dottorato in Biologia Applicata (XXVIII ciclo)



Dottoranda: Roberta Iacono *Roberta Iacono*

Tutor: Dr. Marco Moracci *Marco Moracci*

Coordinatore: Prof. Ezio Ricca

*You can not discover new oceans
until you have the courage to lose sight of the beach*
(Anonymous)

~~~~~Table of contents~~~~~

Keywords	1
List of abbreviations	2
Chapter 1 General Introduction	4
1.1- Carbohydrate Active enZymes	5
1.1.1-CAZy database	6
1.1.2-Glycoside Hydrolases	8
1.1.3-Carbohydrate Esterase	10
1.1.4-Selection of novel enzymes	12
1.2- Extremophiles	14
1.3- Biotechnological applications of thermophilic Glycoside Hydrolase: lignocellulose degrading enzymes	19
1.4- Purpose of thesis	28
1.5- References	30
Chapter 2 Novel thermophilic hemicellulases for the conversion of lignocellulose for second generation biorefineries	40
2.1- General overview	42
2.1.1- Renewable energies: bioethanol	43
2.1.2- First vs second generation bioethanol	45
2.1.3- New Enzymes for second-generation biorefineries	48
2.2- Aims of this study	52
2.3- Experimental Procedures	54
2.4- Results and Discussion	64
2.5- References	91
Chapter 3 Temporal changes of microbial community in Pisciarelli Solfatara	98
3.1- General overview	100
3.1.1- Metagenomic of hyperthermophilic environments	101
3.1.2- Metagenomic analysis of Pisciarelli Solfatara	105
3.2- Purpose of this study	110
3.3- Experimental Procedures	112
3.4- Results and discussion	116
3.5- References	130
Chapter 4 Identification and characterization of the first de-N-acetylase from the hyperthermophilic archaeon <i>Sulfolobus solfataricus</i>	134
4.1- General overview	136
4.1.1- <i>Sulfolobus solfataricus</i>	137

4.1.2- <i>N</i> -acetyl-glucosamine in Sulfolobales	139
4.2- Purpose of work	142
4.3- Experimental Procedures	144
4.4- Results and discussion	153
4.5- References	174
 Publications	 183
Conferences	195
Acknowledgment	197

~~~~~Keywords~~~~~

Glycoside hydrolases
Second generation biorefineries
Hemicellulase
Archaea
Extreme environments
Metagenomic
Sulfolobus solfataricus
N-acetyl-glucosamine
De-N-acetylases

~~~~~List of Abbreviations~~~~~

4MU	4-methylumbelliferyl
4Np-GlcUA	4-Nitrophenyl- α -glucuronide
AA	Auxiliary activities
CAZy	Carbohydrate Active Enzymes database
CAZymes	Carbohydrate Active Enzymes
CE	Carbohydrate Esterase
Cel	Cellobiose
CMC	Carboxyl-methyl-cellulose
DGGE	Denaturing Gradient Gel Electrophoresis
EPS	Exopolysaccharides
FCE	Free cell extract
FSA	Fluorescamine
GalNAc	<i>N</i> -acetyl-galactosamine
GH	Glycoside Hydrolases
Glc	Glucose
GlcN	Glucosamine
GlcNAc	<i>N</i> -acetyl-glucosamine
GlcNAc-1P	<i>N</i> -acetyl-glucosamine 1-phosphate
GlcNAc-6S	<i>N</i> -acetyl-glucosamine 6-sulfate
GPI	Glycosyl Phosphatidyl Inositol
GT	Glycosyl transferase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hex	Hexose
HPAEC-PAD	High- Performance Anion-Exchange chromatography with Pulsed Amperometric Detection
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ManNAc	<i>N</i> -acetyl-mannosamine
mDNA	Metagenomic DNA
MeGlcA	4- <i>O</i> -Methyl-D-glucuronic acid
MGX	4- <i>O</i> -Methyl-D-glucuro-D-xylan
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
PI	Phosphatidyl Inositol
TLC	Thin Layer Chromatography
UDP	Uridine 5'-diphosphate
Xyl	Xylose
Xyl2-5	Xylobiose/triose/tetraose/pentaose

Chapter 1 – General Introduction

Carbohydrates are among the most functionally and structurally diverse molecules in the biological world. In contrast to DNA, RNA, and proteins that are linear polymers containing only one basic type of linkage between monomers, carbohydrates can form branching structures and the multiple monosaccharide building blocks can be linked through various regio- and stereochemistries.

Glycans can accumulate as stores of energy, such as glycogen and starch, or can exploit structural role as chitin, cellulose and hemicelluloses. The latter polysaccharides are the most abundant biopolymers on Earth and can accumulate to such extraordinary abundances because they are difficult to break down: the calculated half-life of the *O*-glycosidic bond in cellulose is 4.7×10^9 years (Wolfenden R et al., 1998). Therefore, carbohydrate active enzymes are among the most efficient biocatalysts known, able to accelerate enormously chemical reactions. The numerous combinatorial possibilities of glycan building blocks explains the variety of enzymatic activities involved in their modification, such as glycosidases, glycosyltransferases and esterases. These enzymes build, modify, and hydrolyze complex carbohydrates and glycoconjugates (glycans bound to proteins or lipids) addressing a large body of biological roles collectively studied under the term of *Glycobiology*.

Here, the discovery of novel enzymes is crucial for both basic and applicative interests. For instance, the efficient deconstruction by a panel of different biocatalysts of plant (hemi)celluloses for fermentations to fuels, chemicals and polymers is a prominent research topic in last few years. In addition, the carbohydrate active enzymes involved in the modification of glycoconjugates acting as defensive barriers, support, store resources, and signaling molecules in Bacteria and Archaea is of great interest (Ridley BL et al., 2001; Carpita NC and Gibeaut DM, 1993). Despite current efforts and technological advances, the discovery and the functional characterization of novel carbohydrate active enzymes still lag behind the constant production of a large mass of (meta)genomic data, and requires a combination of different approaches.

1.1 – Carbohydrate Active enZymes

1.1.1 CAZy Database

The diversity of complex carbohydrates is controlled by a panel of enzymes involved in their assembly (glycosyltransferases, GT) and breakdown (glycoside hydrolases, GH; polysaccharide lyases, PL; carbohydrate esterases, CE), collectively designated as Carbohydrate-Active enZymes (CAZymes).

In 1991, Dr. B. Henrissat proposed a classification of GHs based on protein sequence and structure similarities that correlates with enzyme mechanisms and stereospecificity (Henrissat B, 1991). This classification is still valid nowadays and has been extended to the Classes of GT, CE, and PL. Each family is annotated with information regarding known enzyme activities and includes catalytic and structural features all determined experimentally. Members of the same family share the same protein fold, reaction mechanism, and catalytic residues. In addition, a family can be divided in subfamilies of enzymes with common characteristics, but acting on different substrates (Ferrara MC *et al.*, 2014). Families with the same tertiary structure are grouped into *clans* which are thought to have evolved from a common ancestral gene (Henrissat B and Bairoch A, 1996). In addition to enzymes, CAZy database also includes Carbohydrate-binding Modules (CBMs) that have no enzymatic activity per se, but potentiate the activity of other CAZymes by targeting to and promoting a prolonged interaction with the substrate. CAZy database is currently available at URL www.cazy.org.

Each family encloses enzymatic activities of different substrate specificities (i.e. enzymes with different EC numbers) (Lombard V *et al.*, 2014) demonstrating that the acquisition of novel specificity has been commonplace during evolution. The processes by which a novel substrate specificity was acquired from a common ancestor leave detectable traces in the sequence of contemporary proteins. While the precise cleavage specificity of DNAses, RNAses and proteases is difficult or impossible to derive from their sequence alone, the GH family

classification system allows in some cases the prediction of the category of carbohydrate substrate, in terms of stereospecificity, catalytic mechanism, mode of action, and type of sugar recognized (Cantarel BL *et al.*, 2012). This feature was subsequently noted for the other classes of CAZymes (Campbell JA *et al.*, 1997; Lombard V *et al.*, 2010), greatly expanding the database. To date, CAZy group 135 GH families, 98 GT families, 24 PL and 16 CE (last updated March 2016). Recently, a new CAZy Class has been defined, named Auxiliary Activities (AA), that group polysaccharide lytic monooxygenases (LPMO), catalyzing the oxidative cleavage of cellulose using reducing agents such as ascorbate, gallate and even lignin (Westereng B *et al.*, 2011). The criteria for integration in the AA classification consists in the potential ability to help the original GH, PL and CE enzymes gain access to the carbohydrates encrusted in the plant cell wall. This novel category groups together the families of LPMO and the families of redox enzymes are involved in lignin breakdown (Levasseur A *et al.*, 2013).

CAZy database project inspired Bruce Stone to develop a comprehensive encyclopedia, named CAZypedia (<http://www.cazypedia.org>), whose current curator is Prof. Herry Brumer. Each page of this resource cites primary literature which reports on the first identification of enzyme active site residues, catalytic mechanism, three dimensional structure, or other important publications which define catalysis in a GH family or subfamily. All contributors to CAZypedia, from the Authors to the Board of Curators, are selected experts in the field and sign the respective categories.

The CAZy database and the CAZypedia resources are helpful tools to derive mechanistic informations, structural data and further information related to the families.

1.1.2 – Glycoside Hydrolases

Glycoside hydrolase are enzymes that catalyze the hydrolysis of glycosidic bonds between two or more carbohydrate residues or a carbohydrate unit linked to an aglycon group. This class of enzymes is widely distributed across prokaryotes and eukaryotes (Henrissat B, 1991; Henrissat B and BairochA, 1993; Cantarel BL *et al.*, 2009) and are present in variable copy numbers among organisms. The great variety amongst GHs is explained by the need of selective hydrolysis of glycosidic bonds for energy uptake, cell wall expansion and degradation, and turnover of signalling molecules. The Enzyme Commission (EC) number corresponding to GHs is EC 3.2.1.x where x represents substrate specificity. Some GHs, acting as phosphorylase, are referred with EC number 2.4.1.x.

Over the CAZy classification described above, GHs are classified based on their catalytic mechanisms proposed by Koshland (Koshland DE, 1953) in *inverting* or *retaining*, depending on the stereochemistry at the anomeric center of the product generated. *Inverting* GHs catalyze the hydrolysis of the glycosidic bonds generating a product with anomeric carbon inverted than that in the substrate. Instead, the *retaining* mechanism leads to a retention of the configuration at the anomeric carbon of the substrate after hydrolysis. In *inverting* enzymes catalysis proceeds through a single-displacement mechanism (Fig.1.1.2).

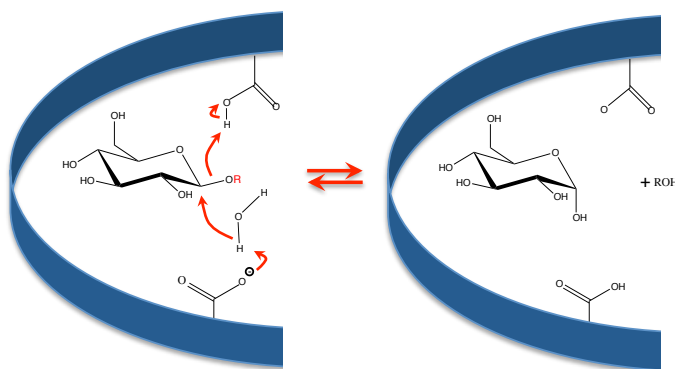


Figure 1.1.2: Catalytic mechanism of an *inverting* β -glucosidase

The catalytic machinery of these enzymes involves two catalytic carboxylic acids usually glutamic or aspartic acids: one functions as general acid, and the other as a general base that assists the nucleophilic attack driven by a water molecule.

By contrast, the reaction by *retaining* GHs proceeds via a double displacement mechanism involving a covalent glycosyl-enzyme intermediate. In the first step of the reaction (glycosylation), the carboxylic residue functions as a general acid catalyst by protonating the glycosidic oxygen while the carboxylate as nucleophile, by attacking the C1 atom of the glycosidic bond and forming a covalent glycosyl-enzyme intermediate. In the second step (de-glycosylation), the carboxylic group previously acting as an acid catalyst, now acts as a general base by deprotonating an incoming water molecule which attacks the anomeric center resulting in the release of the sugar molecule with the same anomeric configuration of the substrate (Figure 1.1.3).

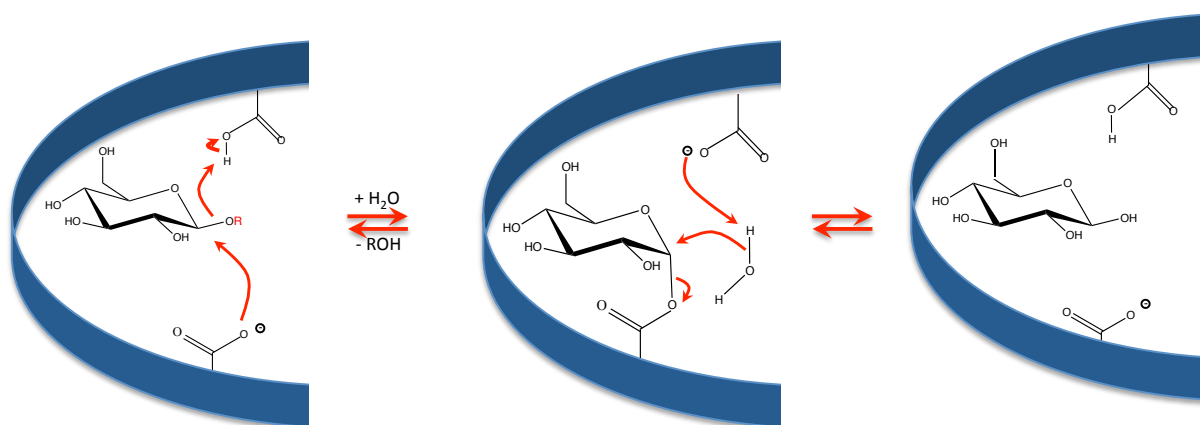


Figure 1.2.2: Catalytic mechanism of a *retaining* β -glucosidase

The GHs can be further classified on the basis of their action mode on long sugar chains: *exo*-glycosidases catalyze the hydrolysis of the glycosidic bond from the reducing or the non-reducing termini of carbohydrate polymers, while the *endo*-glycosidases hydrolyzes the bonds in the middle of a poly- or oligosaccharide (Fig. 1.1.3).

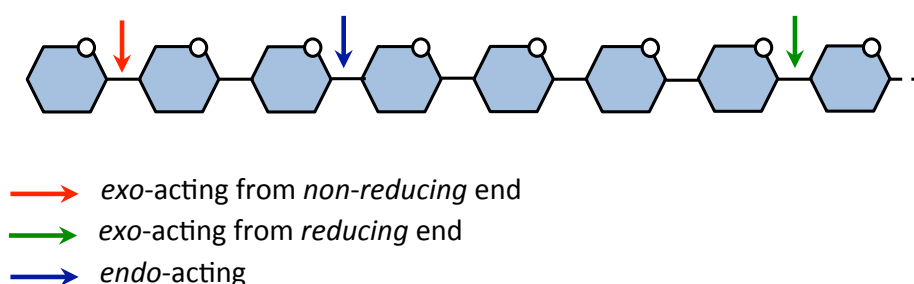


Figure 1.1.3: Mode of action of glycoside hydrolase

1.1.3 – Carbohydrate Esterase

The CAZy classification currently lists 16 different families of carbohydrate esterases/deacetylases. Biologically, these enzymes are involved in the removal

of *O*-ester and *N*-acetyl moieties from carbohydrates; the similarity between sugar and non-sugar esterases, and the promiscuity of their substrate specificity make in many cases the CE classification less insightful and predictive than that of GH, GT and PL.

The majority of *O*-esterases whose structures have been solved display a classical $\beta/\alpha/\beta$ 'serine protease' fold, as revealed by three-dimensional structures of the enzymes from family CE1 (bacterial ferulate esterases), CE5 (acetyl xylan esterases) and CE7 (multifunctional and xylooligosaccharide deacetylases) (Vincent F *et al.*, 2003). The Enzyme Commission (EC) number of *O*-esterase is of the kind EC 3.1.1.x while EC number of de-*N*-acetylase is EC 3.5.1.x, where x represents substrate specificity.

De-*N*-acetylases are classified in families CE4, CE9, CE11, CE14. In CE4 family are grouped de-*N*-acetylases sometimes refereed to as 'NodB homologs' (Davies GJ *et al.*, 2005), which are involved in the deacetylation of peptidoglycan, chitin, rhizobial Nod factors (Blair DE *et al.*, 2005), and *O*-esterase acting on xylan. CE9 are involved in *N*-acetyl-glucosamine-6-phosphate deacetylation (Vincent F *et al.*, 2004) while CE11 zinc-dependent LpxC acts on UDP-3-*O*-acetyl-*N*-acetylglucosamine. CE14 family encloses D-myo-inosityl-2-amino-2-deoxy- α -D-glucopyranoside deacetylase and *NN'*-deacetyl-chitobiose de-*N*-acetylase from Archaea and Bacteria (Lombard V. *et al.*, 2014). Interestingly, the catalytic mechanism of members from this family proceeds via a chemical mechanism in which the carbonyl group of the substrate is polarized by a catalytic metal ion upon binding of the substrate. The side chain of a carboxylate functions as a general base catalyst to activate the metal-water for attack on the carbonyl substrate while the catalytic metal ion and a Tyr residue stabilize the resulting oxyanion tetrahedral intermediate. Finally, the side chain of an His functions as a general acid catalyst to facilitate breakdown of the tetrahedral intermediate (Huang X and Hernick M, 2012) (Figure 1.1.4).

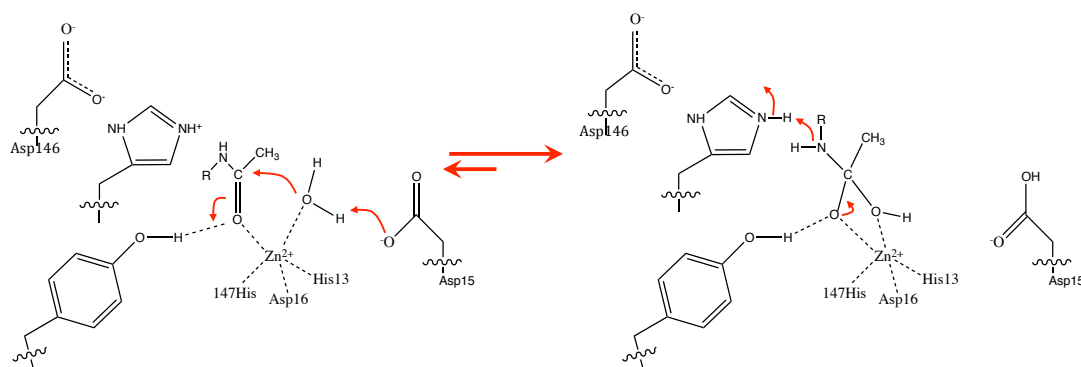


Figure 1.1.4: Catalytic mechanism of de-*N*-acetylase MshB.

1.1.4 - Selection of novel enzymes

The discovery of new extremophilic microorganisms and their enzymes has had a great impact on the field of biocatalysis, since these microorganisms produce unique cazyms that function under conditions in which their mesophilic counterparts would quickly denature, permitting their employment in industrial processes. In this context, the question is how to find the catalysts with the desired properties and robustness. Several options may be considered: the consultation of gene and protein sequence data banks, which includes metagenomic databases including genes from uncultured species, can be performed to dig deep in the diversity of annotated sequences. Under these regards, the CAZy database and the community-driven resource CAZypedia are efficient tools to integrate the current database annotations in an effort of predicting the general function, fold and mechanism. Then, these predictions *in-silico* have to be rigorously tested through detailed experimental characterization of the recombinant enzymes.

Alternatively, a potent tool to discover new GH is the functional screenings on cell. Two basic generic strategies are commonly used: (i) positive selection *in vivo* of growing cells on a medium containing a substrate for the enzyme of interest as the sole source of carbon or as enzyme-reporter assays (e.g. a chromogenic or fluorescent substrates) or (ii) enzymatic assays on cellular

lysates. Using these methods, numerous glycosidases, such as an endo and exo-acting-glycanases, were isolated, (Ferrara *et al.*, 2014; Farrow MF and Arnold FH, 2011)

In this scenario, the work presented here shows how both the approaches of the database inspection and functional screening of cell lysates led to the identification of novel thermophilic carbohydrate active enzymes.

1.2 –Extremophiles

Microorganisms adapted to survive in ecological niches such as high or low temperatures and/or pH, high salt concentrations and high pressure are defined by the anthropocentric point of view as “extremophiles”. The majority of the organisms growing in extreme environments belong to Archaea that, firstly proposed by Carl Woese as the third Domain of life on Earth, are different from Bacteria and Eukarya (Figure 1.2.1) (Woese CR *et al.*, 1990; Woese CR *et al.*, 1991).

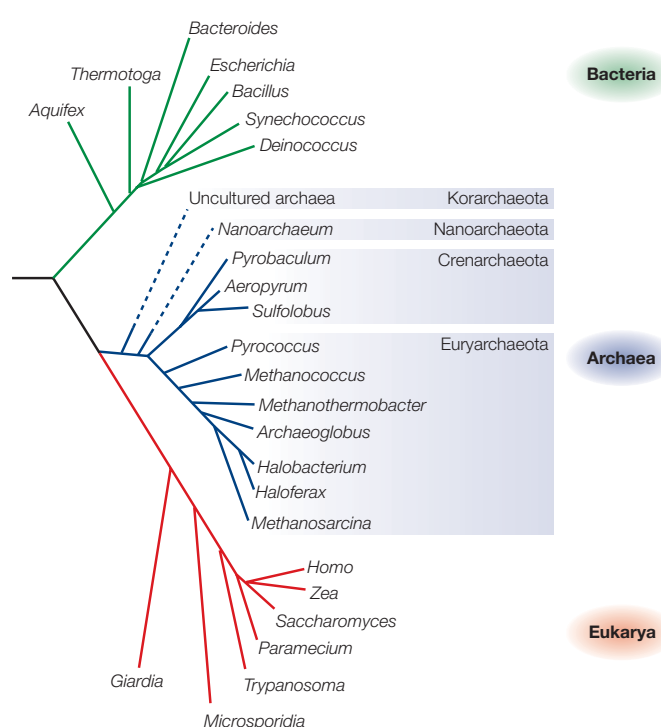


Figure 1.2.1: Phylogenetic tree of life

The Archaea share traits with both Bacteria and Eukarya: their biological information processing systems, such as DNA replication, transcription and translation, is similar under several aspects to those of Eukarya while metabolism is in common with the Bacteria (for instance Entner-Doudoroff pathway) (Table 1.2.1). Moreover, Archaea hold several peculiar traits as they are the only life forms known to carry out methanogenesis to produce biological methane (Cavicchioli R, 2011).

1.2- Extremophiles

Trait	Bacteria	Archaea	Eukarya
Carbon linkage of lipids	Ester	Ether	Ester
Phosphate backbone of lipids	Glycerol-3-phosphate	Glycerol-1-phosphate	Glycerol-3-phosphate
Metabolism	Bacterial	Bacterial-like	Eukaryotic
Core transcription apparatus	Bacterial	Eukaryotic-like	Eukaryotic
Translation elongation factors	Bacterial	Eukaryotic-like	Eukaryotic
Nucleus	No	No	Yes
Organelles	No	No	Yes
Methanogenesis	No	Yes	No
Pathogens	Yes	No	Yes

Table 1.4.1: Main different traits between Bacteria, Archaea and Eukarya

Archaea have traditionally been classified on the basis of their lifestyles as methanogens, extreme halophiles (haloarchaea), and thermoacidophiles. Methanogens are found in anaerobic marine and freshwater environments and in the gastrointestinal tracts of animals, where they participate in the conversion of organic matter by utilizing the metabolic products of Bacteria (for example, CO₂, H₂, acetate and formate) and converting them into methane (CH₄). Haloarchaea reside in hypersaline environments (such as salterns, lakes and the Dead Sea), where they grow as heterotrophs, often in association with phototrophic algae. Thermophiles (including hyperthermophiles, which colonize environments with temperatures above 80°C and (hyper)thermoacidophiles, which grow at pH 2-3) colonize volcanic terrestrial environments and deep-sea hydrothermal vents, growing aerobically or anaerobically as heterotrophs or autotrophs, and often deriving energy by sulphur oxidation or reduction (Antranikian G, 2005).

More recently, a more rigorous classification of Archaea based upon SSU rRNA gene sequence analysis, led to five distinct phyla:

Euryarchaeota. The name derives from the greek term ευρυς, that means “broad” and reflects the phenotypic heterogeneity of this phylum. Included among the Euryarchaeota are various orders of methanogens (Methanobacteriales, Methanococcales, Methanomicrobiales,

Methanosarcinales, and Methanopyrales), as well as extreme halophiles (Halobacteriales), sulfate reducers (Archaeoglobales), and various thermophiles (Thermococcales and Thermoplasmatales).

Crenarchaeota. This phylum was first described by Woese (Woese CR *et al.*, 1990). It shows a limited phenotypic heterogeneity with respect to Euryarchaeota. All pure cultures of this phylum are thermophiles and hyperthermophiles.

Some crenarchaeal members have the ability to grow autotrophically using a recently identified 3-hydroxypropionate/4-hydroxybutyrate carbon dioxide assimilation pathway (Berg IA *et al.*, 2007). Crenarchaea of the orders Sulfolobales and Desulfurococcales utilize a unique Cdv cell division machinery that is related to eukaryotic membrane remodeling systems. Phylogenetically, Crenarchaeota represents a sister group to the clade comprising Thaumarchaeota and the candidate phylum 'Aigarchaeota' (Guy L and Ettema TJ, 2011).

Korarchaeota. This phylum is geographically restricted to terrestrial and marine thermal environments. The only genome sequence currently available for this candidate phylum is *Candidatus* Korarchaeum cryptofilum, which was enriched in lab (Elkins JG *et al.*, 2008). The predicted gene functions suggest that the organism relies on a simple mode of peptide fermentation as source of carbon and energy (Elkins JG *et al.*, 2008)

Nanoarchaeota. The unique representative of this phylum is *Nanoarchaeum equitans*, a tiny symbiont of the Crenarchaeota *Ignicoccus hospitalis*. The phylogenetic analysis of concatenated ribosomal proteins suggested that this species diverged before the separation of the Crenarchaeota and the Euryarchaeota, thus defining Nanoarchaeota as a very ancient phylum (Petitjean C *et al.*, 2014).

Thaumarchaeota. This phylum (Brochier-Armanet C *et al.*, 2008) groups

abundant chemolithoautotrophic ammonia-oxidizers that play an important role in biogeochemical cycles in both aquatic and terrestrial environments, such as the nitrogen and carbon cycles. Thaumarchaeota are also referred to as 'mesophilic or low-temperature Crenarchaeota' and were first discovered in marine environments (DeLong EF, 1992; Fuhrman JA *et al.* 1992) but are now also known to reside in terrestrial habitats (Spang A *et al.*, 2010).

Archaea have featured prominently in hypotheses for the origin of Eukaryotes, as Eukaryotes and Archaea represented sister lineages in Woese's 'universal tree' (Woese CR *et al.*, 1990). Recent phylogenetic analyses of universal protein data sets have provided increasing support to models indicating that eukaryotes emerge as sister to or from within the archaeal 'TACK' superphylum (Guy L and Ettema TJ, 2011), a clade originally comprising the archaeal phyla Thaumarchaeota, Aigarchaeota, Crenarchaeota and Korarchaeota (Cox CJ *et al.*, 2008; Foster PG *et al.*, 2009; Guy L *et al.*, 2014; Lasek-Nesselquist E and Gogarten JP, 2013; Williams TA *et al.*, 2012). This hypothesis has been supported by the recent sequencing of the genome of Lokiarchaeota, a novel, deeply rooting clade of the archaeal TACK superphylum, which forms a monophyletic group with eukaryotes in phylogenomic analyses of universal proteins. (Spang A *et al.*, 2015).

1.3 – Biotechnological applications of thermophilic Glycoside Hydrolase: lignocellulose degrading enzymes

Following the discovery of the first thermophilic microorganisms, their enzymes showing extreme thermal stability coupled to their resistance to extremes of pH and high concentrations of detergents or organic solvents, gained interest as biocatalyst with great industrial potential (Bruins ME *et al.*, 2001; Demirjian DC *et al.*, 2001; Niehaus FC *et al.*, 1999). Running industrial processes at elevated temperatures has numerous advantages: (i) reduction of the risk of contamination as compared to low temperature, (ii) decreased viscosity and increased diffusion coefficient, leading to a higher bioavailability and solubility of organic compounds, and thereby providing a more efficient bioremediation (Becker P *et al.*, 1997), (iii) improved hydrolysis of hardly degradable or non-degradable polymers.

Thermostable GHs are widely used in industrial processes, for instance the pectinases are widely used in extraction, clarification, and removal of pectin in fruit juices, in maceration of vegetables to produce pastes and purées, and in wine-making (Tu T *et al.*, 2014). GHs are “friendly” catalysts also in other industrial processes. For example, xylanases constitute a major group of industrial enzymes with significant application in paper industry: the hydrolysis of xylan at high temperature facilitates release of lignin from paper pulp and reduces the level of usage of chlorine as the bleaching agent (Shallom D and Shoham Y, 2003). α -Galactosidase are employed in the food and feed industries to eliminate the α -D-galactosides (mainly raffinose and stachyose), the main anti-nutritional factor in legume seeds, and to improve digestibility (Wang C *et al.*, 2016). However, current efforts are directed to optimize the role of thermostable GHs in the emerging field of renewable energies. GHs are widely employed in starch conversion and lignocellulosic degradation, for the production of first and second generation bioethanol, respectively. Starch is composed by α -glucose units, linked by α -1,4- and α -1,6- glycosidic bond, thereby forming the two high-molecular-weight components amylose (20%)

and amylopectin (80%). The starch hydrolysis (saccharification) by α -amylase and α -glucosidase releases the glucose monomers into the solution. The subsequent or simultaneous fermentation with *Saccharomyces cerevisiae* strains, convert the sugar monomers into ethanol and carbon dioxide (Lennartsson PR *et al.*, 2014).

In comparison, the lignocellulosic material requires a vast panel of enzymes for its degradation because of their complex structure. The composition and the enzymes involving in its hydrolysis are discussed in detail below.

Lignocellulose degradation: (hemi)cellulose composition

The structure of lignocellulosic material has been selected in nature to resist the microbial degradation and to maintain the erected structure enabling the plants to capture the sunlight. The recalcitrance of these structural materials is based on a complex polymeric structure with variable proportions of cellulose, hemicellulose and lignin ([Figure 1.3.1](#)).

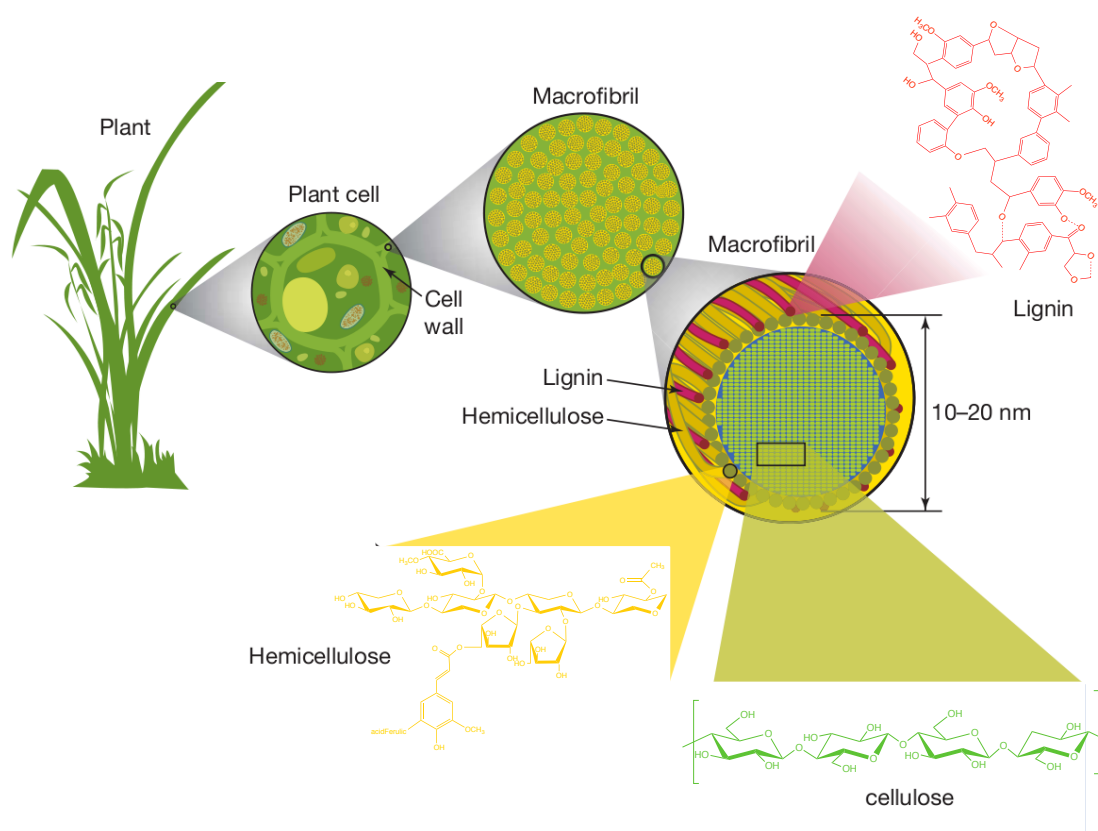


Figure 1.3.1: Structure of plant cell walls

Cellulose is a simple linearly linked polymer of β -1,4-D-glucose units; the polysaccharide chains bind one to another by hydrogen bonding excluding the water and making difficult the penetration and activity of degradative enzymes. Hemicellulose polymers have higher structural diversity and are usually classified according to the predominant monosaccharide building block forming the main polymeric chain, such as (glurono)xylan, (gluco-, galacto)mannan, xyloglucans, etc. or because of different etymology, such as chitin from $\chi\iota\tau\omega\nu$ (coating). These polysaccharides can also be branched with monomers of D-xylose, D-galactose, L-arabinose, D-glucuronic acid, and other monosaccharides, raising their complexity and diversity of structures which often, depend on the plant species, degree of maturation, origin of plant part, etc. (Scheller HV and

Ulvskov P, 2010). Main plant hemicelluloses are shown in Figure 1.3.2. The most important biological role of hemicelluloses is their contribution to strengthening the cell wall by interaction with cellulose and lignin (Scheller HV and Ulvskov P, 2010). This latter is an unusually complex biopolymer lacking defined primary structure and formed by aromatic alcohols, known as monolignols, and is considered a main defence barrier against enzymatic degradation of biomasses (Bhalla A *et al.*, 2012).

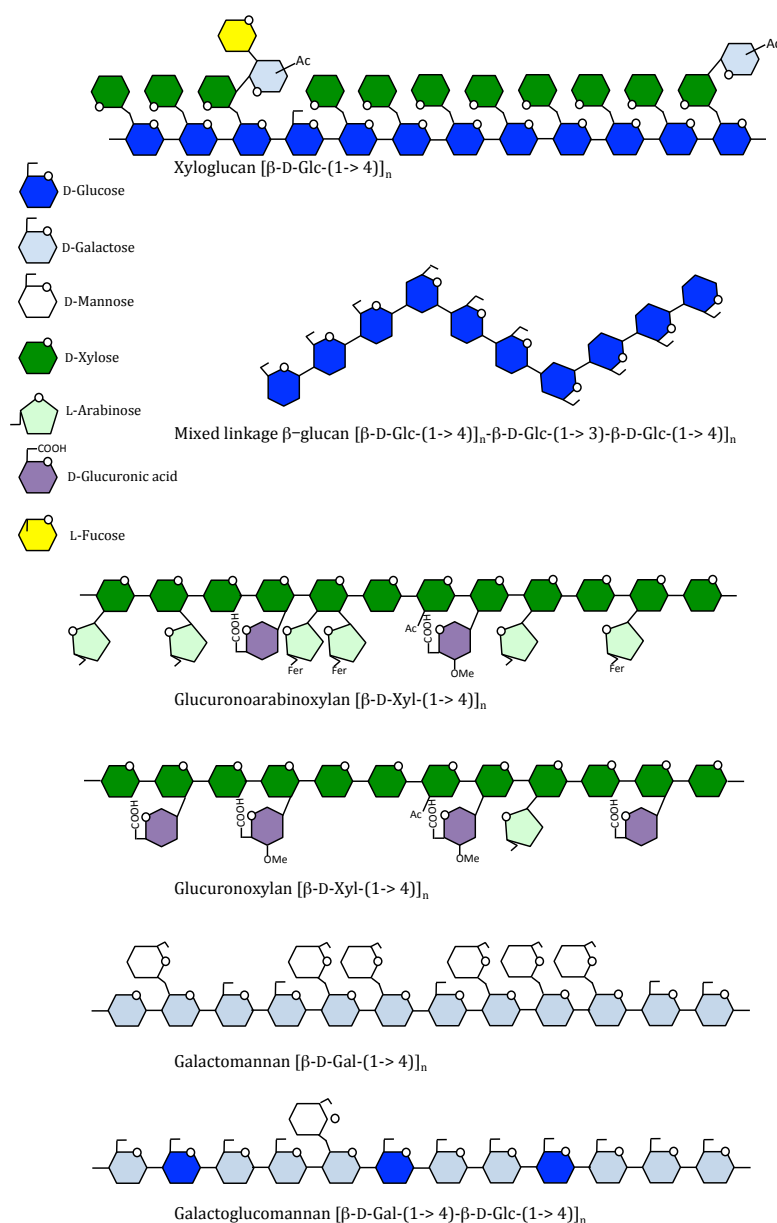


Figure 1.3.2: Different types of hemicelluloses

Lignocellulose degradation: cellulase and hemicellulase

Cellulases are enzymes that catalyze the hydrolysis and subsequent depolymerization of cellulose. It has been reported that three enzymatic activities, endoglucanase, cellobiohydrolase and β -glucosidase work synergistically to completely and efficiently hydrolyze cellulose (Table 1.3.1).

Enzyme degrading cellulose	EC number	CAZy classification
Endo-glucanase	3.2.1.4	GH5, GH6, GH7, GH8, GH9, GH12, GH44, GH51, GH59, GH74, GH124
Cellobiohydrolase (CBH)	3.2.1.91	GH5, GH6, GH7, GH8, GH9, GH48
β -glucosidase	3.2.1.21	GH1, GH3, GH5, GH9, GH30, GH46
Copper-dependent lytic polysaccharide monooxygenases (LPMOs)	-	AA9, AA10, AA11

Table 1.3.1: Enzymes degrading cellulose classification

Endoglucanase hydrolyzes cellulose, producing oligosaccharides, and traces of cellobiose and glucose. Cellobiohydrolase (CBH) hydrolyzes β -1,4-D-glucosidic linkages in cellulose oligosaccharides, releasing cellobiose from either the non-reducing or the reducing ends while β -glucosidase hydrolyzes cellobiose to glucose. Often, endoglucanases contain carbohydrate-binding modules (CBM), that improve the efficiency in hydrolyzing crystalline cellulose by the catalytic cellulase domain (Yeoman CJ *et al.*, 2010). Endoglucanases are widespread among GH families and catalyze the hydrolysis through both *retaining* and *inverting* mechanisms according to the belonging family. The majority of thermostable endoglucanases belong to GH12 family. (Ando S *et al.*, 2002; Kim JO *et al.*, 2000; Park C *et al.*, 2001). As well as CBHs catalyze the hydrolysis of glycosidic bonds with retention of configuration except the members of families GH6 and GH7 (Khademi S *et al.*, 2002; Rouvinen J *et al.*, 1990).

β -Glucosidases may be divided into three groups on the basis of substrate specificity: aryl- β -glucosidases, cellobiases, and broad-specificity β -glucosidases. Aryl- β -glucosidases exhibit an extreme preference toward hydrolysis of β -glucosides bound to compounds different from sugars (e. g. alcohols, lipids, etc.), whereas cellobiases hydrolyze cello-oligosaccharides only (including cellobiose). Members of the third group, termed broad-specificity β -glucosidases, show significant activity on both substrate types and represent the most commonly observed group in cellulolytic microbes (Bhatia Y *et al.*, 2002). Except for GH9, all β -Glucosidases hydrolyze their substrate through a *retaining* mechanism.

In addition, the cellulose degradation can be improved by the copper-dependent lytic polysaccharide monooxygenases (LPMOs) (families AA9, AA10 and AA11) that cleavage the cellulose through the oxidation of the glycosidic bonds (Beeson WT *et al.*, 2012).

In nature, fungi produce more cellulases than bacteria, however, cellulases produced by the latter have more applications in biorefineries as they encounter less glucose feedback inhibition (Acharya S and Chaudhary A, 2012). In addition, bacterial cellulases have very high activities against crystalline celluloses like cotton or Avicel, are more thermostable and are more active at alkaline pH if compared to fungal cellulases (Bon EPS *et al.*, 2002; Macedo JMB *et al.*, 1999; Zerbini JE *et al.*, 1999). In Bacteria, cellulolytic activities have been found in *Bacillus* strains (Mawadza C *et al.*, 2000), with optimal activity at 60°C, and in the thermophilic anaerobic bacterium *Clostridium thermocellum* (Mori Y, 1992). In particular, *C. thermocellum* degrade the cellulose by cellulosome, an enzymatic multicomplex composed by different types of cellulose-degrading enzymes and hemicellulolytic enzymes assembled on the structural scaffoldin subunits through strong non- covalent protein-protein interactions between the docking

modules (dockerin) and complementary modules (cohesins) (Dashtban M *et al.* 2009). Highly thermostable cellulases acting at 95°C have been reported from *Thermotoga maritima* MSB8 (Bronnenmeier K *et al.*, 1995). An endocellulase, with the ability to hydrolyze microcrystalline cellulose and maximal activity at 85–95°C, was isolated from the extremely thermophilic bacterium *Anaerocellum thermophilum* (Zverlov V *et al.*, 1998). Thermostable cellulases of archaeal origin showing optimal activity at 102–105°C have been isolated from *Pyrococcus furiosus* and *Pyrococcus horikoshii* (Kengen S *et al.*, 1993; Ando S *et al.*, 2002). The Crenarchaea *Sulfolobus solfataricus* MT4, *Sulfolobus acidocaldarius*, and *Sulfolobus shibatae* have also been shown to produce significant amount of β -glucosidases (Grogan W, 1991).

Among hemicelluloses, xylan is one of the most common in the feedstocks used in biorefineries. The complete enzymatic hydrolysis of xylan requires the synergistic action of several enzymatic activities, summarized in Table 1.3.2 (Subramaniyan and Prema, 2002).

Hemicellulase	EC number	CAZy classification
Endo- β -xylanase	3.2.1.8 3.2.1.32	GH5, GH8, GH10, GH11, GH26, GH30, GH43, GH51
α -arabinofuranosidase	3.2.1.55	GH5, GH3, GH43, GH51, GH54, GH62
α -glucuronidase	3.2.1.139	GH4, GH67, GH115
β -xylosidase	3.2.1.37	GH3, GH39, GH43, GH52, GH54
Acetyl-xylan esterase	3.1.1.72	CE1, CE2, CE3, CE4, CE5, CE6, CE7, CE12
Feruloyl esterase	3.1.1.73	CE1

Table 1.3.2: Enzymes degrading xylan classification

Endo- β -1,4-xylanase hydrolyze the xylan backbone producing xylooligosaccharides, decorated with α -1,2 and α -1,3 linked glucuronate and arabinosides residues, respectively; each removed by α -glucuronidases and α -L-arabinofuranosidases. Depending on maturation step of plants, α -arabinosides

are acetylated or esterified with ferulic acid; in this case the acetylxyylan esterases and feruloyl esterases actions are required. A more detailed description of these enzymes is shown in Chapter 2 (Paragraph 2.1.3).

Thermostable xylanases have been isolated from a variety of sources including terrestrial and marine solfataric fields, thermal springs, hot pools and self-heating decaying organic debris (Singh S *et al.*, 2003; Harris GW *et al.*, 1997; Vieille C and Zeikus GJ 2001; Cannio R *et al.*, 2004). The majority of these belong to families 10 and 11. Family 10 xylanases have been isolated from various hyperthermophilic and thermophilic Bacteria including *Thermotoga sp.* (Winterhalter C *et al.*, 1995; Zverlov V *et al.*, 1996), *B. stearothermophilus* T-6 (Khasin A *et al.*, 1993) *Bacillus sp.* NG-27 (Lo Leggio L *et al.*, 1999), *Bacillus sp.* N16-5 (Ma Y *et al.*, 2004) *B. halodurans* (Mamo G *et al.*, 2006) *B. firmus* (Chang P *et al.*, 2004) *Caldicellulosiruptor sp.* (Luthi E *et al.*, 1990) *Clostridium thermocellum* (Herbers K *et al.*, 1995), *Rhodothermus marinus* (Abou-Hachem M *et al.*, 2003), and *Thermoascus aurantiacus* (Lo Leggio L *et al.*, 1999, Natesh R *et al.*, 2003). A family 10 xylanase isolated from *Thermotoga sp.* Strain FjSS3-B.1 is one of the most thermostable xylanases reported with an optimum temperature for activity at 105 °C and pH 5.5 and a half life of 90 minutes at 95 °C (Simpson HD *et al.*, 1991).

1.4- Purpose of thesis

Carbohydrates play an important role in a variety of biological and industrial process. The discovery of novel thermophilic CAZymes can help us to understand the functional involvement of this class of enzymes in various biological phenomena and how we can exploit their “eco-friendly” activities in numerous industrial processes.

In this framework, my thesis work is dedicated to the study of thermophilic CAZymes and can be subdivided in three principal sections, Chapters 2, 3 and 4, respectively. Chapter 2 is focused on the identification of novel thermophilic glycosidases selected in sequence database. Since sustainable biorefineries require numerous biocatalysts, our research topic was addressed at the identification of new thermostable hemicellulases that can be used in the production of second generation bioethanol. The search of novel thermostable enzymes led us to sequence metagenomic DNA isolated from a solfataric field. In this context, we explored the changes of microbiome colonizing this environment in relation to geothermal variations, as described in Chapter 3. This study allows to increase our knowledge on habitats with high selective pressure. In Chapter 4, we focused on the selection of novel cazymes from crenarchaeon *S. solfataricus* P2 by functional screening. This study led us to achieve additional data concerning the set of cazymes present in *S. solfataricus* allowing to identification of a novel unclassified enzyme.

1.5- References

- Abou-Hachem M, Olsson F, Nordberg Karlsson E (2003) Probing the stability of the modular family 10 xylanase from *Rhodothermus marinus*. *Extremophiles* 7: 483–491
- Acharya S, Chaudhary A. (2012). Bioprospecting thermophiles for cellulase production: a review. *Braz J Microbiol.* 2012 Jul;43(3):844-56.
- Ando S, Ishida H, Kosugi Y, Ishikawa K (2002). Hyperthermostable endoglucanase from *Pyrococcus horikoshii*. *Appl. Environ. Microbiol.* 2002;1:430–433.
- Antranikian G (2005) .Extreme Environments as a Resource for Microorganisms and Novel Biocatalysts. *Adv Biochem Engin/Biotechnol* 2005 96: 219–262
- Becker P, Abu-Reesh I, Markossian S, Antranikian G, Märkl H. (1997). Determination of the kinetic parameters during continuous cultivation of the lipase-producing thermophile *Bacillus* sp. IHI-91 on olive oil. *Appl Microbiol Biotechnol.* 1997 Aug;48(2):184-90.
- Beeson WT, Phillips CM, Cate JH, Marletta MA (2012). Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc.* 2012 Jan 18;134(2):890-2.
- Berg IA, Kockelkorn D, Buckel W, Fuchs G. (2007) A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* 318, 1782–1786
- Bhalla A, Bansal N, Kumar S, Bischoff KM, Sani RK (2013) Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes, *Bioresour. Technol.* 128 2013 751–759.
- Bhatia Y, Mishra S, Bisaria VS. (2002) Microbial beta-glucosidases: cloning, properties, and applications. *Crit Rev Biotechnol.* 2002;22(4):375-407. Review.
- Blair DE, Schuttelkopf AW, MacRae JA, van Aalten DMF, Brochier-Armanet C. et al. (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245–252

- Bronnenmeier K, Kern A, Libel W, Staudenbauer W (1995) Purification of *Thermotoga maritima* enzymes for the degradation of cellulose materials. *Appl. Environ. Microbiol.* 1995;61:1399–1407.
- Bruins ME, Janssen AE, Boom RM. (2001) Thermozyms and their applications: a review of recent literature and patents. *Appl Biochem Biotechnol.* 2001 Feb;90(2):155-86.
- Campbell JA, Davies GJ, Bulone V and Henrissat B (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.*, 326, 929–939.
- Cannio R, Di Prizito N, Rossi M, Morana A (2004) A xylan-degrading strain of *Sulfolobus solfataricus*: isolation and characterization of the xylanase activity. *Extremophiles* 8: 117–124
- Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 2009 Jan;37(Database issue):D233-8.
- Cantarel BL, Lombard V and Henrissat B (2012) Complex carbohydrate utilization by the healthy human microbiome. *PLoS One*, 7, e28742
- Carpita NC, Gibeaut DM (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth, *Plant J.* 3 1993 1–30.
- Cavicchioli R. (2011) Archaea-timeline of the third domain. *Nat Rev Microbiol.* 2011 Jan;9(1):51-61.
- Chang P, Tsai WS, Tsai CL, Tseng MJ (2004) Cloning and characterization of two thermostable xylanases from an alkaliphilic *Bacillus firmus*. *Biochem Biophys Res Commun* 319: 1017–1025
- Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM. (2008) The archaeobacterial origin of eukaryotes. *Proc. Natl Acad. Sci. USA* 105, 20356–20361 2008
- Dashtban M, Schraft H, Qin W. (2009) Fungal bioconversion of lignocellulosic residues; opportunities & perspectives. *Int J Biol Sci.* 2009 Sep 4;5(6):578-95.

- Davies GJ, Gloster TM, Henrissat B. (2005) Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr Opin Struct Biol.* 2005 Dec;15(6):637-45. Review.
- DeLong EF. (1992) Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. U.S.A.* 89, 5685–5689
- Demirjian DC, Morís-Varas F, Cassidy CS. (2001) Enzymes from extremophiles. *Curr Opin Chem Biol.* 2001 Apr;5(2):144-51.
- Elkins JG, Podar M, Graham DE, Makarova KS, Wolf Y, Randau L, Hedlund BP, Brochier-Armanet C, Kunin V, Anderson I, Lapidus A, Goltsman E, Barry K, Koonin EV, Hugenholtz P, Kyrpides N, Wanner G, Richardson P, Keller M, Stetter KO. (2008) A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8102–8107
- Farrow MF, Arnold FH. (2011) High throughput screening of fungal endoglucanase activity in *Escherichia coli*. *J Vis Exp.* 2011 Aug 13;(54). pii: 2942.
- Ferrara MC, Cobucci-Ponzano B, Carpentieri A, Henrissat B, Rossi M, Amoresano A, Moracci M. (2014) The identification and molecular characterization of the first archaeal bifunctional exo- β -glucosidase/N-acetyl- β -glucosaminidase demonstrate that family GH116 is made of three functionally distinct subfamilies. *Biochim Biophys Acta.* 2014 Jan;1840(1):367-77.
- Foster PG, Cox CJ, Embley TM. (2009) The primary divisions of life: a phylogenomic approach employing composition-heterogeneous methods. *Phil. Trans. R. Soc. Lond. B* 364, 2197–2207 2009.
- Fuhrman JA, McCallum K, Davis AA. (1992) Novel major archaeobacterial group from marine plankton. *Nature* 356, 148–149
- Grogan W. (1991) Evidence that β -galactosidase of *Sulfolobus solfataricus* is only one of several activities of a thermostable β -D-glycosidase. *Appl. Environ. Microbiol.* 1991;57:1644–1649.
- Guy L, Ettema TJ. (2011) The archaeal 'TACK' superphylum and the origin of eukaryotes. *Trends Microbiol.* 2011 Dec;19(12):580-7.

- Guy L, Saw JH, Ettema TJ (2014) The archaeal legacy of eukaryotes: a phylogenomic perspective. *Cold Spring Harb. Perspect. Biol.* 6, a016022 2014.
- Harris GW1, Pickersgill RW, Connerton I, Debeire P, Touzel JP, Breton C, Pérez S. (1997) Structural basis of the properties of an industrially relevant thermophilic xylanase. *Proteins* 29: 77–86
- Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 1991 280, 390-316
- Henrissat B, Bairoch A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem J.* 1996 Jun 1;316 (Pt 2):695-6.
- Herbers K, Wilke I, Sonnewald U (1995) A thermostable xylanase from *Clostridium thermocellum* expressed at high-levels in the apo-plast of transgenic tobacco has no detrimental effects and is easily purified, *Bio/technology* 13:63–66
- Huang X, Hernick M. (2012) Examination of mechanism of N-acetyl-1-D-myo-inositol-2-amino-2-deoxy- α -D-glucopyranoside deacetylase (MshB) reveals unexpected role for dynamic tyrosine. *J Biol Chem.* 2012 Mar 23;287(13):10424-34.
- Kengen S, Luesink E, Stams A, Zehnder A. (1993) Purification and characterization of an extremely thermostable β -glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur. J. Biochem.* 1993;1:305–312.
- Khademi S, Guarino LA, Watanabe H, Tokuda G, and Meyer EF. (2002) Structure of an endoglucanase from termite, *Nasutitermes takasagoensis*. *Acta Crystallogr. D. Biol. Crystallogr.* 58, 653–659.
- Khasin A, Alchanati I, Shoham Y (1993) Purification and characterization of a thermostable xylanase from *Bacillus stearothermophilus* T-6. *Appl Environ Microbiol* 59: 1725–1730
- Kim JO, Park SR, Lim WJ, Ryu SK, Kim MK, An CL, Cho SJ, Park YW, Kim JH, Yun HD (2000). Cloning and characterization of thermostable endoglucanase (Cel8Y) from the hyperthermophilic *Aquifex aeolicus* VF5. *Biochem. Biophys. Res. Commun.* 279, 420–426.

- Koshland DE. 1953. Stereochemistry and the mechanism of enzymatic. *Biological Reviews* 28:416–436.
- Lasek-Nesselquist E, and Gogarten JP. (2013) The effects of model choice and mitigating bias on the ribosomal tree of life. *Mol. Phylogenet. Evol.* 69, 17–38 2013.
- Lennartsson PR, Erlandsson P, Taherzadeh MJ. (2014) Integration of the first and second generation bioethanol processes and the importance of by-products. *Bioresour Technol.* 2014 Aug;165:3-8.
- Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels.* 2013 Mar 21;6(1):41.
- Lo Leggio L, Kalogiannis S, Bhat MK, Pickersgill RW (1999) High resolution structure and sequence of *T. aurantiacus* xylanase I: implications for the evolution of thermostability in family 10 xylanases and enzymes with (beta)alpha-barrel architecture. *Proteins* 36: 295–306
- Lombard V, Bernard T, Rancurel C, Brumer H, Coutinho PM, Henrissat B. (2010) A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem. J.*, 432, 437–444.
- Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013 *Nucleic Acids Res.* 2014 Jan;42(Database issue):D490-5.
- Luthi E, Bhana Jasmat N, Bergquist PL (1990) Xylanase from the extremely thermophilic bacterium “*Caldocellum saccharolyticum*”: Overexpression of the gene in *Escherichia coli* and characterisation of the gene product. *Appl Environ Microbiol* 56: 2677–2683
- Ma Y, Xue Y, Dou Y Xu Z, Tao W, Zhou P. (2004) Characterization and gene cloning of a novel beta-mannanase from alkaliphilic *Bacillus* sp N16–5. *Extremophiles* 8: 447–454
- Macedo JM, Gottschalk LM, Bon EP. (1999) Lignin peroxidase and protease production by *Streptomyces viridosporus* T7A in the presence of calcium carbonate. Nutritional and regulatory carbon sources. *Appl Biochem Biotechnol.* 1999 Spring;77-79:735-44.

- Mamo G, Delgado O, Martinez A, Mattiasson B, Kaul RJ (2006) Cloning, sequencing analysis and expression of a gene encoding an endoxylanase from *Bacillus halodurans* S7. *Mol Biotechnol* 33: 149–159
- Mawadza C, Hatti-Kaul R, Zvaunya R, Mattiasson B. (2000) Purification and characterization of cellulases produced by two *Bacillus* strains. *J. Biotechnol.* 2000;83:177–187.
- Mori Y. (1992) Comparison of the cellulolytic systems of *Clostridium thermocellum* YMA and JW20. *Biotechnol. Lett.* 1992;14:131–136.
- Natesh R, Manikandan K, Bhanumoorthy P, Viswamitra MA, Ramakumar S (2003) Thermostable xylanase from *Thermoascus aurantiacus* at ultrahigh resolution (0.89 Å) at 100 K and atomic resolution (1.11 Å) at 293 K refined anisotropically to small-molecule accuracy. *Acta Crystallogr D* 59: 105–117
- Niehaus F, Bertoldo C, Kähler M, Antranikian G. (1999) Extremophiles as a source of novel enzymes for industrial application. *Appl Microbiol Biotechnol.* 1999 Jun;51(6):711-29.
- Park C, Kawaguchi T, Sumitani J, Arai M. (2001). Purification and characterization of cellulases (CBH I and EGL 1) produced by thermophilic microorganism M23. *Appl. Biol. Sci.* 7, 27–35. *Streptomyces* sp.
- Petitjean C, Deschamps P, López-García P, Moreira D (2014) Rooting the domain archaea by phylogenomic analysis supports the foundation of the new kingdom Proteoarchaeota *Genome Biol Evol.* 2014 Dec 19;7(1):191-204.
- Ridley BL, O'Neill MA, Mohnen D. (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling, *Phytochemistry* 57 2001 929–967.
- Rouvinen J, Bergfors T, Teeri T, Knowles JK, Jones TA. (1990) Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science* 249, 380–386.
- Scheller HV and Ulvskov P (2010) Hemicelluloses *Annu. Rev. Plant Biol.* 2010. 61:263–89

- Shallom D, Shoham Y. (2003) Microbial hemicellulases. *Curr Opin Microbiol.* 2003 Jun;6(3):219-28.
- Simpson HD, Haufler UR, Daniel RM (1991) An extremely thermostable xylanase from the thermophilic eubacterium *Thermotoga*. *Biochem J* 15: 413–417
- Singh S, Madlala AM, Prior BA (2003) *Thermomyces lanuginosus*: properties of strains and their hemicellulases. *FEMS Microbiol Rev* 27: 3–16
- Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, Ettema TJ. (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature*. 2015 May 14;521(7551):173-9.
- Spang, A. Hatzenpichler R, Brochier-Armanet C, Rattei T, Tischler P, Spieck E, Streit W, Stahl DA, Wagner M, Schleper C. (2010) Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends Microbiol.* 18, 331–340
- Subramaniyan S and Prema P. (2002) Biotechnology of microbial xylanases: enzymology, molecular biology, and application. *Crit Rev Biotechnol.* 2002;22(1):33-64.
- Tu T, Meng K, Huang H, Luo H, Bai Y, Ma R, Su X, Shi P, Yang P, Wang Y, Yao B. (2014) Molecular Characterization of a Thermophilic Endopolygalacturonase from *Thielavia arenaria* XZ7 with High Catalytic Efficiency and Application Potential in the Food and Feed Industries [dx.doi.org/10.1021/jf504239h](https://doi.org/10.1021/jf504239h) | *J. Agric. Food Chem.* 2014, 62, 12686–12694
- Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* 65: 1–43
- Vincent F, Charnock SJ, Verschueren KHG, Turkenburg JP, Scott DJ, Offen WA, Roberts S, Pell G, Gilbert HJ, Davies GJ, Brannigan JA. (2003) Multifunctional xylooligosaccharide/cephalosporin C deacetylase revealed by the hexameric structure of the *Bacillus subtilis* enzyme at 1.9Å^o resolution. *J Mol Biol* 2003, 330:593-606

- Vincent F, Yates D, Garman E, Davies GJ, Brannigan JA. (2004) The 3-D structure of the N-acetylglucosamine-6-phosphate deacetylase, NagA, from *Bacillus subtilis*: a member of the urease superfamily. *J Biol Chem* 2004, 279:2809-2816.
- Wang C, Wang H, Ma R, Shi P, Niu C, Luo H, Yang P and Yao B. (2016) *J Biosci Bioeng*. 2016 Biochemical characterization of a novel thermophilic α -galactosidase from *Talaromyces leycettanus* JCM12802 with significant transglycosylation activity *Jan*;121(1):7-12.
- Westereng B, Ishida T, Vaaje-Kolstad G, Wu M, Eijsink VG, Igarashi K, Samejima M, Ståhlberg J, Horn SJ, Sandgren M. The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS One*. 2011;6(11):e27807.
- Williams TA, Foster PG, Nye TM, Cox CJ, Embley TM (2012) A congruent phylogenomic signal places eukaryotes within the Archaea. *Proc. R. Soc. Lond. B* 279, 4870–4879 2012.
- Winterhalter C, Heinrich P, Candussio A, Wich G, Liebl W. (1995) Identification of a novel cellulose-binding domain within the multidomain 120 kDa xylanase XynA of the hyperthermophilic bacterium *Thermotoga maritima*. *Mol Microbiol* 15: 431–444.
- Woese CR, Achenbach L, Rouviere P, Mandelco L. (1991) Archaeal phylogeny: reexamination of the phylogenetic position of *Archaeoglobus fulgidus* in light of certain composition-induced artifacts. *Syst Appl Microbiol*. 1991;14(4):364-71.
- Woese CR, Kandler O, Wheelis ML. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A*. 1990 Jun;87(12):4576-9.
- Wolfenden R, Lu X, and Young G. (1998) Spontaneous Hydrolysis of Glycosides *J. Am. Chem. Soc.* 1998, 120, 6814-6815.
- Yeoman CJ, Han Y, Dodd D, Schroeder CM, Mackie RI, Cann IK. (2010) Thermostable enzymes as biocatalysts in the biofuel industry. *Adv Appl Microbiol*. 2010;70:1-55. doi: 10.1016/S0065-2164(10)70001-0.

- Zerbini JE, Oliveira EMM, Bon EPS (1999) Lignin peroxidase production by *Streptomyces viridosporus* T7A: Nitrogen nutrition optimization using glucose as carbon source. *Appl. Biochem. and Biotechnol.* 1999:77–79.
- Zverlov V, Piotukh K, Dakhova O, Velikodvorskaya G, Borriss R (1996) The multidomain xylanase A of the hyperthermophilic bacterium *Thermotoga neapolitana* is extremely thermoresistant. *Appl Microbiol Biotechnol* 45: 245–247.
- Zverlov V., Riedel K., Bronnenmeier K. (1998) Properties and gene structure of a bifunctional cellulytic enzyme (CelA) from the extreme thermophile *Anaerocellum thermophilum* with separate glycosyl hydrolase family 9 and 48 catalytic domains. *Microbiology.* 1998;144:457–465.

Chapter 2~
Novel thermophilic hemicellulases for the
conversion of lignocellulose for second
generation biorefineries

Chapter 2 of this thesis describes the identification, biochemical characterization, and exploitation of three hemicellulases on a pretreated material from an energy crop. In particular, here the work is focused on the identification of novel highly thermostable endo- β -xylanase and exo- α -glucuronidases, involved synergically in the deconstruction of the hemicellulose. This section begins with a brief introduction on renewable energies (paragraph 2.1.1) and the differences between first and second generation bioethanol (paragraphs 2.1.2). The paragraph 2.1.3 describes the endo- β -xylanases and α -glucuronidases. The purpose of this work is shown in paragraph 2.2, while the experimental part is described in paragraph 2.3. Results and discussion are described in paragraph 2.4. At last, the paragraph 2.5 reports the literature references included in this chapter.

2.1- Introduction

2.1.1 – Renewable energies: bioethanol

Energy is becoming a big issue in our society because of the growing awareness of the finite resources of fossil liquid fuels and the noticeable climate changes resulting from its massive use. The accumulation of huge amounts of CO₂ derived from burning of fossil fuels causes significant and negative effects on the global environment. The atmospheric CO₂ concentrations increased by almost 100 ppm from their pre-industrial level, reaching 400 ppm for the first time in 2015 as a consequence of human activities (Figure 2.1.1). The International Panel on Climate Change (IPCC) predicts that by the year 2100, the atmosphere may contain up to 570 ppm. It is thus crucial to take effective measures to reduce CO₂ emissions and use the available CO₂.

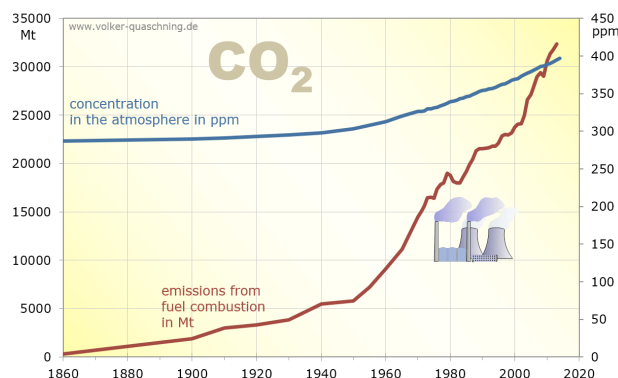


Figure 2.1.1: CO₂ emissions from 1860 to 2015 (Source: International Energy Agency).

Looking for alternative and renewable energies could be the only solution for the sustainable development of our society (Du XL *et al.*, 2015). One essential field is the energy conversion from non-fossil fuel sources and the storage of the gained energy. Among the different type of renewable energy, such as solar, eolic, nuclear and electric energy, bioethanol is a major player in the challenge to reduce the use of fossil fuels. The success of bioethanol is due to several advantages: (i) bioethanol can be blended with fossil fuel, such as E10 (10% ethanol and 90 % gasoline) that can be used in all motors reducing furthermore

the amount of high-octane additive (Aditiya HB *et al.*, 2014). (ii) The use of ethanol-blended fuels such as E85 (85% ethanol and 15% gasoline) can reduce the net emissions of greenhouse gases by as much as 37.1% (Aditiya HB *et al.*, 2014). Moreover, (iii) an other important advantage in the bioethanol use is that the feedstocks can be domestically-produced: this makes the countries without crude oil resources able to produce energy, gaining some economic freedom.

However, the disadvantages of the bioethanol consist in the much lower energy content than petrol: indeed burning 1 liter of ethanol gives 34% less energy than burning the same amount of petrol. In other words, bioethanol is not as efficient as petroleum (Aditiya HB *et al.*, 2014). In addition, the high octane number of bioethanol (105) makes difficult to burn it. Therefore, most biofuels retain at least a small amount of petrol, such as E85. Finally, the main problem is the bioethanol produced by food feedstock (first generation bioethanol) that requires a large amount of arable land for the growing of energy crops. This can lead the farmers to convert the agricultural lands at the energy crop production, since the bioethanol is a product to higher added-value. Consequently, not enough food could be available to feed the growing population and in turn cause a significant increase in food prices. Moreover, the destruction of some natural habitats and the competition of food vs fuel may also occur (Lombardi M *et al.*, 2014). The European Commission has explored possible options to mitigate worries from indirect land use change (ILUC) by elaborating the Staff Working Document — SWD(2012) (European Commission, 2012). In particular, this draft proposal, issued on the 17th October 2012, defines a precautionary threshold of first generation biofuels (initially 5%, then adjusted to 7% in 2015 with the Directive 2015/1513) in the transportation market (European Commission, 2012; European Commission, 2015). Its main purpose is to enhance the production of second-generation biofuels, e.g. from waste feedstock and residues, which can usually ensure larger GHG savings (Rana R *et al.*, 2015).

2.1.2- First vs second generation bioethanol

Currently, all industrial scale production of ethanol belongs to the first generation of biofuels. The largest ethanol producing countries are USA and Brazil. The sugar-based ethanol is predominantly produced in Brazil from sugarcanes while the starch-based ethanol is generally from corn but also from grains, and is dominated by the US followed by China, Canada, France, Germany, and Sweden (REN 21, 2012). The first step of the ethanol production from grains (dry mills) is the milling of the substrate and subsequent liquefaction of the starch. The liquefaction is followed by the hydrolysis or saccharification, which releases the sugar (glucose) monomers into the solution. During the subsequent, or simultaneous, fermentation with yeast, the sugar monomers are converted into ethanol and carbon dioxide. Usually, an ethanol concentration of ca. 10% (w/v) is obtained at the end of the fermentation. The fermentation liquid is distilled to separate and purify the ethanol; thus, even if the process only provides a low rate of return, it is rather convenient in terms of cost of the feedstock and price of the products. However, the use of potential human food as feedstock has led to considerable ethical discussions, normally referred to as the “food vs fuel” debate, with widely diverse and strongly polarized views. The supply of the feedstock can also become a potential limiting factor compared with the potential demand. Moreover, its use do not allow to gain fuel and economic freedom, as it depends on specific feedstock produced only in some countries (Lennartsson PR *et al.*, 2014). By contrast, second generation ethanol utilizes different types of “not food” lignocellulosic materials as substrate. From this material alone, Kim and Dale reported that the global ethanol production can potentially produce about 442 billion litres, which is approximately 16 times greater compared to the current global production (Kim S and Dale B, 2014).

The production of second generation bioethanol is summerized in [Figure 2.1.2](#)

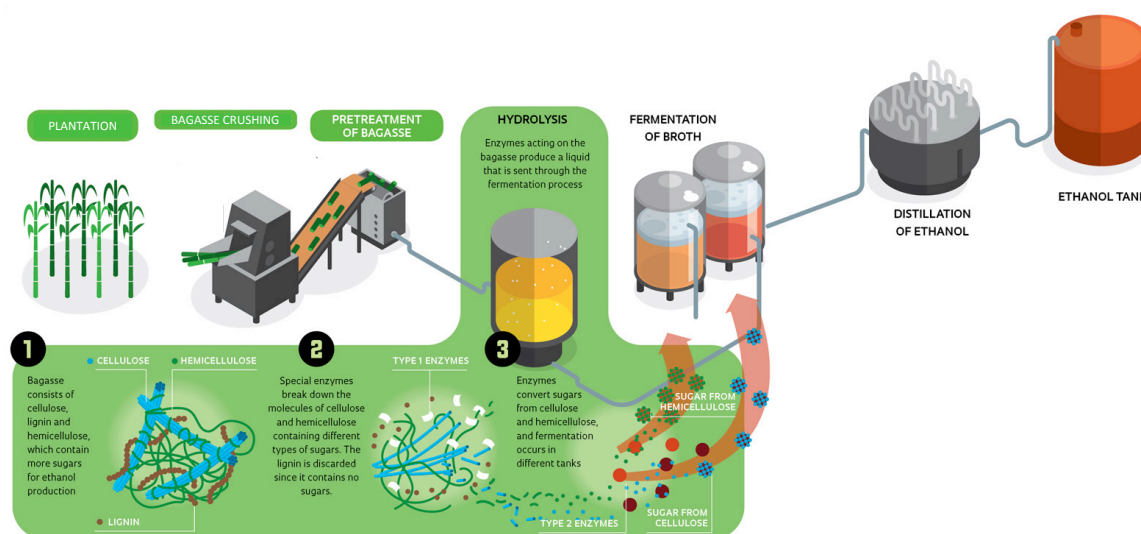


Figure 2.1.2: Production process of second generation bioethanol.

Currently, only negligible amounts of second generation bioethanol are produced in several plants around the world that work industrially, but are not yet commercially feasible. Second generation ethanol processes have technically no issues as far as feedstock supply, as 7–18 billion tons/year of lignocellulosic biomass is available for human exploitation (Lin Y and Tanaka S, 2006). Corn stover, wheat straw, and switch grass are examples of lignocellulosic materials (Mosier N *et al.*, 2005). Generally, the type of plant source of lignocellulosic material (the so called “energy crop”) is chosen depending on specific properties: it should be tolerant to a wide range of environmental stresses and cultivated on marginal, degraded, or contaminated lands, thus reducing competition with food crops. In addition, it should be autochthonous to reduce the costs and the CO₂ emissions due to the transport.

In European context, the energy balance of first-generation bioethanol is quite low compared to second generation bioethanol: in terms of energetic yields, measured as Energy Returned On Energy Invested (EROEI), is 1 for bioethanol from corn while is 10 for bioethanol from lignocellulose (Østergård H *et al.*, 2012). However, the process is currently limited by technical and economic

challenges that, although connected, can be identified into three types (Cheng JJ and Timilsina GR, 2011). The first technical challenge is caused by the recalcitrance of the biomass and thus the need of harsh pretreatments of the feedstock. These pretreatments (steam-explosion, hydrothermal, and pretreatment with acid or alkali), often result in the formation of compounds inhibiting the following fermentation. The second challenge is in the production of efficient enzymes to hydrolyze the (hemi)cellulose, at a cost competitive to the first generation enzymes hydrolyzing starch. Although great efforts have been accomplished by the enzyme manufacturers, improvements are still necessary. Thirdly, sufficiently high concentrations of ethanol (e.g. 4– 4.5% w/v) have to be reached in order to reduce the cost of distillation and wastewater treatment (Lennartsson PR *et al.*, 2014). Another issue can be the fermenting microorganism of choice, since common *S. cerevisiae* exploited in first-generation bioethanol is unable to utilize pentoses. A plethora of examples of genetic manipulation to overcome this hamper exists in the literature (Madhavan A *et al.*, 2012). However, legal issues and consumer opinions regarding the use of genetically modified organisms, especially in Europe, are often overlooked.

To overcome one of the major obstacle represented by enzymatic hydrolysis for production at industrial scale, the highly efficient thermophilic cellulase and hemicellulase attracted much attention for reducing the cost of biofuels production.

2.1.3- New Enzymes for second-generation biorefineries

Ethanol production from lignocellulosic biomass includes pretreatment of biomass, enzymatic hydrolysis of (hemi)cellulose, fermentation of hexose/pentose sugars, and recovery of ethanol. Intensive efforts have been made in recent years to develop efficient technologies for the pretreatment, for enhanced cellulose/hemicellulose enzymatic saccharification, and for the fermentation of C6 and C5 sugars. Enzymatic hydrolysis is an ideal approach for degrading (hemi)cellulose into fermentable sugars because it does not present corrosion problems in the reactors and result in negligible by-products formation with high sugar yields (Peng X *et al.*, 2015). However, the process depends on optimized conditions for maximal efficiency (temperature, time, pH, enzyme and substrate concentrations) and may suffer from end-product inhibition and biomass structural restraints (Vimala Rodhe A *et al.*, 2011). To overcome these problems, the use of thermophilic enzyme could be an interesting solution. For the efficient hydrolysis of different types of lignocellulosic materials, enzyme mixtures have to be customized. In particular, the development of these *cocktails* requires better knowledge about the specific activities involved, in order to optimize the process. In addition to cellulose, whose degradation has been described in Paragraph 1.4, the most abundant hemicellulose in nature is xylan which contains mainly β -D-xylopyranosyl residues linked by β -1,4-glycosidic bonds (Uday US *et al.*, 2016). Due to its structural heterogeneity, complete degradation requires the synergistic action of different enzymatic activities (Figure 2.3.1). Among these, endoxylanases are the major enzymes responsible for the hydrolysis of xylan (Uday US *et al.*, 2016).

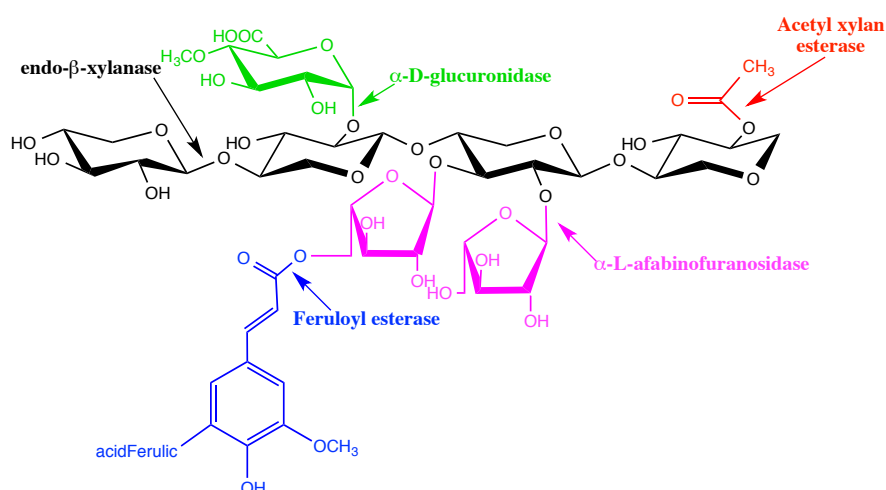


Figure 2.3.1: Enzymatic activities involved in xylan degradation

In CAZy classification, endo-1,4- β -xylanase activity can be found in several families (Lombart V *et al.*, 2014) that differ in their physicochemical properties, structure and substrate specificities. Like GH10, GH5 and GH30 xylanases are structurally similar to each other, showing a $(\beta/\alpha)_8$ -fold, but while GH5 xylanases are more specific to arabinoxylan (Bhardwaj A *et al.*, 2014; Correia MAS *et al.*, 2011), GH30 xylanases are appendage-dependent and need free 4-*O*-methyl-D-glucuronosyl (MeGlcA) residues as side chain to be active (St. John FJ *et al.*, 2006; Vrřanská M *et al.*, 2007). Instead, GH10 xylanases are much more versatile and have a broad substrate specificity: they are highly active on short xylo-oligosaccharides, capable of hydrolysing aryl β -glycosides of xylobiose and xylotri-ose, but are not active on cellulose (Biely P *et al.*, 1997; Gilkes NR *et al.*, 1991). GH11 xylanases display a β -jelly-roll structure and are active on aryl- β -xylo-oligosaccharides but not on aryl- β -cello-oligosaccharides (Biely P *et al.*, 1997; Krenzel U and Dijkstra BW, 1996). Enzymes from GH8 display a $(\alpha/\alpha)_6$ -fold and, in contrast to GH5, GH10 and GH11 xylanases, follow an *inverting* mechanism (Collins T *et al.*, 2002; Van Petegem F *et al.*, 2003).

α -Glucuronidases are less studied hemicellulases and debranch the 4-*O*-methyl-glucuronic acid linked on xylan backbone. They belong to families GH4, GH67,

and GH115 (Lombard V *et al.*, 2014). Family GH4 contains several enzymatic activities and do not hydrolyze 4-*O*-methyl- α -glucuronoxylan or its oligosaccharidic fragments (Suresh C *et al.*, 2003). These enzymes require NAD⁺ and a divalent metal ion and in some instances reducing environments for catalytic activity (Thompson J *et al.*, 1998). GH115 enzymes include only α -glucuronidases and remove glucuronic acid from both the terminal and the internal regions of xylooligosaccharides and xylans (Tenkanen M *et al.*, 2000; Ryabova O *et al.*, 2009). GH67 remove selectively the 4-*O*-methyl- glucuronide- α -1,2 bound to the non-reducing end xylose of short oligosaccharides of glucurono-xylans (Nurizzo D *et al.*, 2002; Golan G *et al.*, 2004), but do not remove glucuronic acid from internal regions of xylan (Ruile P *et al.*, 1997). In addition, they do not display activity on synthetic substrate as 4-nitrophenyl- α -glucuronide.

β -Xylosidases are *exo*-type glycosidases that hydrolyze short xylooligomers into single xylose units, and are found in families GH3, GH39, GH43, GH52, and GH54 (Lombard V *et al.*, 2014). Because of the structural similarity between D-xylopyranose and L-arabinofuranose, compounds with both these monosaccharides, such as 4-nitrophenyl-xyloside and 4-nitrophenyl-arabinoside, can be substrates of bifunctional xylosidase-arabinosidase enzymes, which are found mainly in families GH3, GH43, and GH54. When these enzymes are tested on natural substrates, most of them show only the release of one type of sugar and are thus from a biological and applied perspective monofunctional. The only true bifunctional AX degrading enzymes, which can release both xylose and arabinose from natural substrates, are plant hydrolases from *Arabidopsis thaliana*, *Hordeum vulgare* and *Medicago sativa* (Lagaert S *et al.*, 2014). GH39 is a relatively small family and the β -xylosidase of this family display much more activity on xylotriose and long xylooligosacchaides while on xylobiose produce a large amounts of transglycosylation products (Lee YE and

Zeikus JG. 1993).

Synergism of these different types of hemicellulases offers interesting opportunities for improving the efficiency of bioconverting native lignocellulosic biomass to fermentable sugars. The study of these enzymes from extremely thermophilic microorganisms allows to take advantage of their adaptability to pretreatment conditions and new opportunities in the production of renewable energies.

2.2- Aim of this study

With the increase in global warming and energy consumption and expected impending shortages of crude oil, there is a considerable and immediate interest in developing alternative energy sources. Bioethanol from lignocellulosic feedstock provides a means to reduce the dependence on fossil fuels as well as to reduce global emissions of greenhouse gases. However, this process needs to rely on consistent and robust enzymes able to promote lignocellulose efficiently and consistently.

On the basis of these observations, this part of my PhD project was aimed to identify novel thermostable hemicellulases. The selection strategy has based on the identification of annotated genes in CAZy database from thermophilic organisms with putative activities of interest. My work was focused on the biochemical characterization of the recombinant form of these enzymes that allowed to reveal their catalytic features. Then, the combined action of selected enzymes has been tested on xylans and on pretreated biomass. The synergic action of these hemicellulases has allowed to develop a novel promising thermostable enzymatic *cocktail*.

2.3- Experimental Procedures

Bacterial Strain

E. coli strains used in this work are TOP10: F- *mcrA* Δ (*mrr-hsdRMS-mcrBC*) Φ 80*lacZ* Δ M15 Δ *lacX74 recA1 araD139* Δ (*ara**leu*)7697 *galU galK rpsL* (StrR) *endA1 nupG* and BL21 star: (DE3) .F⁻ *ompT hsdSB* (rB⁻ mB⁻) *gal dcm rne131*(DE3) Invitrogen.

Culture media

LB (Luria-Bertani Broth) (1 liter): 10 g NaCl, 5 g yeast extract, 10 g tryptone

Reagents

All commercially available substrates (4Np- α -GlUA, 4Np- β -Xyl, 2Np- β -Cel, Avicel, CMC, Aldouronic acids, beechwood xylan, 4-*O*-methyl-glucuronoxylan, xyloligosaccharides, cellotriose and cellobiose) were purchased from Sigma-Aldrich, Carbosynth and Megazyme. The biomass of *A. donax* used in this study derived from a steam-explosion pre-treatment described in patent WO2010113129A2.

Cloning of recombinant GH10-XA, GH67-GA and GH67-GC

The Aaci2328 gene, coding GH10-XA xylanase, and the Aaci_0060 gene, codifying GH67-GA α -glucuronidase were amplified by PCR from the genome of *A. acidocaldarius* ATCC27009 while the Csac_2689 gene, encoding GH67-GC, was amplified by PCR from the genome of *C. saccharolyticum* using the synthetic oligonucleotides reported in [Table 2.3.1](#).

2.3- Experimental Procedures

GH10-XA	Aaci2328FWD-5'-CACCATGACGGATCAAGCGCCGT-3'
	Aaci2328REV-5' - GTTTTGGGCGAGGCGCACCAC-3'
GH67-GA	Aaci0060FWD-5' -CACCCCTTGACGAACATCCCTGA-3'
	Aaci0060REV-5' - CGGGTAGATGTGGAGCCC-3'
GH67-GC	Csac2689FWD-5' -CACCATGGAACACGTCAAACAAAAA-3'
	Csac2689REV-5' -TGGATATATAAGTCTTCCTTTTTCATC-3'

Table 2.3.1: PCR primers of GH10-XA, GH67-GA and GH67-GC

The amplification reactions of Aaci_2328 and Aaci_0060 genes were performed with PfuUltra II Fusion HS DNA Polymerase (Stratagene) while the amplification reaction of Csac_2689 was performed with Platinum Taq High Fidelity (Invitrogen). The programs were shown schematically in Figure 2.3.1:

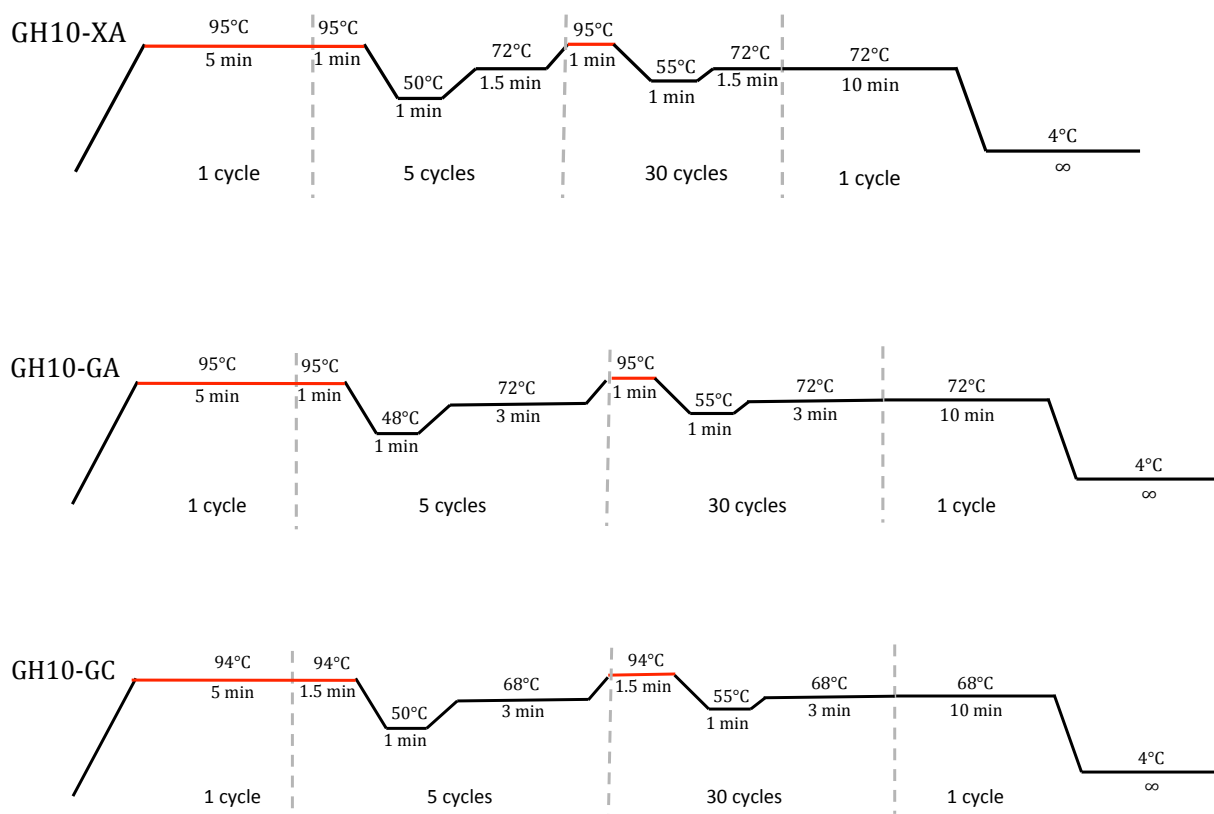


Figure 2.3.1: Scheme of PCR programs of GH10-XA and the GH67s

The DNA fragments obtained were verified by electrophoresis on 1% agarose gel, purified by PCR Kleen spin columns (Biorad) and cloned in the expression vector pET101/D-TOPO (Invitrogen), obtaining the recombinant plasmids pET101/D-TOPO-Aaci 2328, pET101/D-TOPO-Aaci 0060 and pET101/D-TOPO-Csac_2689. The genes are under the control of an isopropyl-1-thio- β -D-galactopyranoside (IPTG) inducible T7 RNA polymerase promoter and the C-terminal of the protein was fused to V5 epitope and 6xHis tag. For the cloning, 100 ng of PCR products Aaci_2328 and Csac_2689 and 50 ng of Aaci_0060 were incubated in salt solution with 1 μ l vector (15-20 ng) at room temperature for 30 minutes. The TOPO cloning reaction was used to transform One Shot TOP10 chemically competent cells according to the protocol of the manufacturer (see below for more details). Positive clones were selected through PCR colony and the absence of mutations was verified by sequencing.

The recombinant plasmids pET101/D-TOPO/Aaci_2328, pET101/D-TOPO/Aaci_0060 and pET101/D-TOPO/Csac_2689 obtained by TOPO cloning reactions described above were transformed into competent One Shot® TOP10 chemically competent *E. coli* cells provided by the kit. The entire TOPO cloning reaction mixtures have been added into a vial of cells. After incubation on ice for 30 minutes, the cells were subjected to a heat-shock at 42°C for 30 seconds and then immediately transferred to ice and supplemented with 250 μ L SOC medium. The cells were incubated at 37°C for 1 hour; then 200 μ L of culture were spread on LB-agar plates containing 50 μ g mL⁻¹ ampicillin and incubated at 37°C for 16 hours.

Expression trials of GH10-XA, GH67-GA and GH67-GC

To determine the better yield of the recombinants GH10-XA, GH67-GA and GH67-GC, expression trials were performed in *E. coli* strain BL21 star (DE3). *E. coli* cells transformed were used to prepare 100 mL culture. The cells were

incubated at 37 °C and when reached an A_{600nm} of 0.6 optical densities the culture was divided in different aliquots for each condition (Figure 2.3.2):

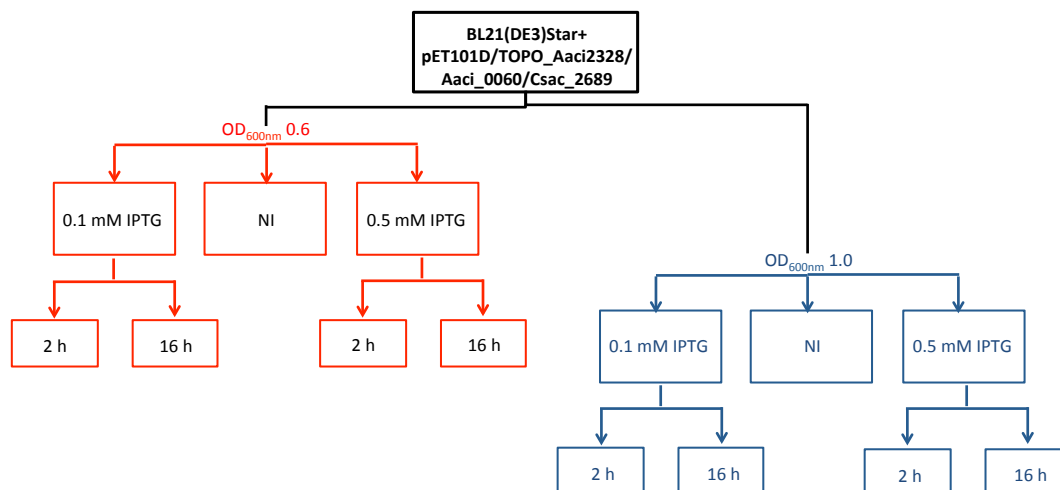


Figure 2.3.2: Dygram of expression trials. NI: not induced.

The cultures were harvested by centrifugation at 5,000 x g for 2 minutes and resuspended in loading buffer 1X. Then, cells were loaded on SDS-PAGE (10%) to evaluate the protein expression.

Purification of GH10-XA, GH67-GA and GH67-GC

E. coli BL21 star (DE3) cells were grown at 37 °C in 2 L of Luria–Bertani (LB) broth supplemented with ampicillin (50 µg mL⁻¹). The induction of GH10-XA recombinant expression was performed by adding 0.1 mM IPTG when the growth was at 1 OD, while GH67-GA and GH67-GC were induced by adding 0.5 mM IPTG when the cultures have reached 1 OD. Growths for expressions were allowed to proceed for 16 h. The cells were harvested by centrifugation at 5000 ×g. The resulting pellets were resuspended in 50 mM sodium phosphate buffer pH 8.0, 300 mM NaCl with a ratio of 5 mL g⁻¹ cells and then were incubated at 37 °C for 1 h with 20 mg of lysozyme (Fluka) and 25 U g⁻¹ cell of Benzonase (Novagen). The cells were lysed by French Press treatment and cell debris were

removed by centrifugation at $10,000 \times g$ for 30 min. The FCEs were loaded on a His Trap FF crude column (GE-Healthcare) equilibrated with Buffer B at a flow rate of 1 mL min^{-1} . After an initial wash-step (20-column volumes) with buffer B, the proteins were eluted with a two-step gradient of imidazole in Buffer B (250 mM imidazole, 20-column volumes; 500 mM imidazole, 20-column volumes).

The protein eluted at 250 mM imidazole. Active fractions were pooled, dialyzed against PBS buffer (buffer phosphate 20 mM, NaCl 150 mM pH 7.3) and then heat-fractionated: or 30 min at 50°C and 60°C for GH10-XA, GH67-GA and GH67-GC and 20 min at 70°C for GH67-GC. The resulting supernatants of GH67-GA and GH67-GC were loaded on a HiLoad 16/60 Superdex column (Amersham Bioscience) equilibrated in PBS buffer. The active fractions were pooled and stored at 4°C . The protein concentration was determined with the Bradford assay.

Molecular masses determination

The molecular mass of GH10-XA, GH67-GA and GH67-GC were determined by gel filtration on a Superdex 200 HR 10/300 FPLC column (GE-Healthcare). Molecular weight markers were Apoferritin (443 kDa), amylase (200 kDa), BSA (66 kDa) and ribonuclease A (13.7 kDa).

Temperature and pH influence

The pH optima were determined by assaying GH10-XA in 50 mM of the indicated buffers at different pHs on 5 mg mL^{-1} of beechwood xylan at 65°C for 1 min. The activity was measured by the Somogyi–Nelson assay (Somogyi M, 1952) estimating the amount of reducing sugars released after 1 min. GH67-GA and –GC were assayed in 100 mM buffers at different pH on 2 mg mL^{-1} of aldouronic acids mixture at 50°C (GH67-GA) and 1 mg mL^{-1} of aldouronic acids mixture at

65 °C (GH67-GC). The temperature profiles were determined in the range of 40-85°C in the assay conditions reported above.

Thermal stabilities were evaluated by incubating the enzymes in PBS buffer, at the indicated temperatures. At intervals, aliquots were withdrawn, transferred in ice and assayed at the same conditions described above. The residual activities were expressed as a percentage of the maximal enzymatic activities measured before the incubation at indicated temperatures.

Enzymatic assays

GH10-XA

The standard assay for GH10-XA activity was performed in 50 mM phosphate buffer pH 6.5 on beechwood xylan at the final concentration of 5 mg mL⁻¹. Typically, in each assay, we used about 1 µg of GH10-XA in the final volume of 0.1 mL. After 1 min of incubation at 65°C, the reaction was blocked in ice. The reaction products were detected by reducing end Somogyi-Nelson assay: in detail, 100 µL of Somogyi-Nelson solution 1 (SN1) were added to GH10-XA reaction. The assay was boiled for 20 min at 100 °C and cooled in ice. After, 100 µL of SN2 were added and the reaction was vortexed and centrifuged for 10 min at 16000 x g. The supernatant was withdrawn and the absorbance was measured at 520 nm at room temperature. In all the assays, spontaneous hydrolysis of the substrate was subtracted by using appropriate blank mixtures without the enzyme, on which the Somogyi-Nelson assay was performed. To measure the specific activity, indicated as U mg⁻¹, a standard curve of glucose (from 0.05 to 0.14 µmoles) was generated for the calculation of enzyme product. One enzymatic unit is defined as the amount of enzyme catalysing the conversion of 1 µmole of reducing ends produced in 1 min at indicated conditions. The enzymatic activity on 4-*O*-methyl-glucuronoxylan (5 mg mL⁻¹), CMC (5 mg mL⁻¹) and Avicel (5 mg mL⁻¹) was evaluated in the same conditions.

The assays on 4Np- β -xyl (8 mM), 4Np- β -glu (5 mM) and 2Np- β -cel (10 mM) were performed in standard conditions. After 1 min of incubation at 65°C, the reactions were blocked in ice by adding 0.8 mL of 1 M sodium carbonate pH 10.2. The absorbances were measured at 420 nm at room temperature and the millimolar extinction coefficients of 4-nitrophenol and 2-nitrophenol were 17.3 and 4.8 mM⁻¹cm⁻¹, respectively. All kinetic data were calculated as the average of at least two experiments and were plotted and refined with the program GraphPad Prism. Kinetic constants of GH10-XA were measured at standard conditions by using concentrations of substrate ranging between 0.04 and 25 mg mL⁻¹ for beechwood xylan, MGX, 4Np- β -xyl and 2Np- β -Cel.

The activity on xyloligosaccharides (1 mg mL⁻¹), cellobiose (10 mM) and cellotriose (10 mM) and was evaluated by Thin Layer Chromatography. TLC analysis of the xylo and gluco-oligosaccharides was carried out by using N-butanol/methanol/water (50:25:25 v/v) as eluent and detected by exposure to 4% α -naphthol in 10% sulphuric acid in ethanol followed by charring. The transglycosylation activity was evaluated by incubating GH10- XA (3 μ g) at standard conditions with 30 mM 4Np- β -Xyl and 8 mM 2Np- β -Cel in a final volume of 0.2 mL. At time intervals (10 min, 60 min and 16 h) aliquots of the reaction were withdrawn and analyzed on TLC, by using ethyl acetate/methanol/water (70:20:10 v/v) as eluent.

The mode of action of GH10-XA was analyzed on 2Np- β -Cel (8 mM) at standard conditions. A time course was carried up to 16 h and the reaction products were analyzed by High- Performance Anion-Exchange chromatography with Pulsed Amperometric Detection (HPAEC-PAD) equipped with a PA200 column (Dionex, USA). HPAEC-PAD analyses of the reactions were performed at a flow-rate of 0.5 mL min⁻¹ in isocratic 20 mM NaOH for 20 min.

GH67-GA and GH67-GC

The activity of GH67-GA was tested on aldouronic acids (2 mg mL^{-1}) in standard conditions: 100 mM buffer phosphate pH 6.5 at 50°C for 1 min with $17 \text{ }\mu\text{g}$ of enzyme. The standard assay of GH67-GC was performed with $10 \text{ }\mu\text{g}$ of enzyme on 1 mg mL^{-1} aldouronic acids in 100 mM buffer phosphate pH 6.0 at 65°C . The assays on aldouronic acids were analyzed by Somogyi-Nelson assay, as described above. GH67-GC was also tested on 0.7 mM 4Np-GlcUA at standard conditions. At 405 nm the molar extinction coefficients of 4-nitrophenol was $9340 \text{ M}^{-1} \text{ cm}^{-1}$. In all the assays, identical mixtures containing all the reagents without the enzymes were prepared as blank. To analyze the activity of GH67-GC on aldouronic acids, aliquots of the reaction were analyzed by TLC, as described above, by using n-butanol/methanol/water (50:25:25 v/v) as eluent and detected by exposure to 4% α -naphthol in 10% sulphuric acid in ethanol followed by charring. The reactions on aldouronic acids were also analyzed by HPAEC-PAD by using the following program: flow 0.4 mL min^{-1} , 35°C , isocratic 100 mM NaOH, [segment 1] 0–25 min up to 125 mM CH_3COONa , [segment 2] 25–31.5 min, up to 450 mM CH_3COONa , [segment 3] up to 35 min, 450 mM CH_3COONa .

Kinetic constants of GH67-GA and GH67-GC on aldouronic acids were measured at standard conditions by using the substrate ranges 0.05–2.5 and 0.02–1 mg mL^{-1} , respectively. Kinetic constants of GH67-GC on 4Np-GluA by using concentrations of substrate ranging between 0.3 and 7 mM. In all the assays, one unit of enzyme activity was defined as the amount of μmoles of produced reducing end by 1 mg of enzyme at the conditions described. Spontaneous hydrolysis of the substrates was subtracted by using appropriate blank mixtures without enzyme. All kinetic data were calculated as the average of at least two experiments and were plotted and refined with the program Prism 5.0 (GraphPad Software, San Diego, California, USA).

Simoultaneous reaction of GH10-XA, GH67-GA, GH67-GC and GH3-XT

To test the effect of simultaneous action of the two enzymes, GH10-XA (0.26 μM) and GH67-GA (0.11 μM) were incubated in 100 mM sodium phosphate buffer pH 6.5, at 50 ° C in the presence of MGX (5 mg mL⁻¹) for 16 h. In addition, the synergical action of GH10-XA and GH67-GC (0.09 μM) was analyzed incubating the two enzyme in 100 mM sodium phosphate buffer pH 6.0, at 65 °C in the presence of MGX (5 mg mL⁻¹) and beechwood xylan (15 mg mL⁻¹) for 2 h. 100 μL of each reaction were withdrawn and the amount of reducing ends was measured as described above. A blank mixture with no enzymes and three aliquots with GH10-XA, GH67-GC and GH67-GA alone were used as controls. The reaction products of GH10-XA and GH67-GC were analyzed by HPAEC-PAD by using the program for aldouronic acids described above. To improve the hydrolysis, GH10-XA and GH67-GC were tested on MGX (5 mg mL⁻¹) and beechwood xylan (15 mg mL⁻¹) in presece of GH3-XT (0.08 μM) for 2 h and the amount of reducing ends were measured as reported above. The activities of GH10-XA, GH67-GC and GH3-XT were tested on the steam-explosion pre-treated *A. donax*. The biomass was diluted 3-fold with water and dialysed (cut-off 1 kDa) against water. GH10-XA (5.1 μM), GH67-GC (0.32 μM) and GH3-XT (0.76 μM) were incubated in 50 mM sodium phosphate buffer, pH 6.0, at 65 ° C in the presence of 0.25 mL of biomass prepared as described above. The reaction products were analyzed by HPAEC-PAD on a CarboPac PA200 Analytical (2 × 205 mm) by using the following program: flow 0.25 mL min⁻¹, 35 °C, [segment 1] 0–20 min isocratic 20 mM NaOH, [segment 2] 20–40 min up to 1 M CH₃ COONa, [segment 3] up to 50 min 1 M CH₃ COONa. Fucose (0.5 nmol) was added as internal standard.

2.4- Results and discussion

Rising global temperature, worsening air quality and drastic declining of fossil fuel reserve are the inevitable phenomena from the disorganized energy management. Second generation bioethanol production captured the attention of many researchers and scientists for a better path of fuel sustainability. Second generation production utilizes the non-edible lignocellulosic and starchy materials from agricultural and forestry biomasses (energy crops) and it becomes one fascinating solution for the problem food-versus-fuel, for global fuel demand and environmental complications (Aditiya HB *et al.*, 2016). Bioethanol production via this route is also expected to be economically preferable in the future for the observable reason of low feedstock cost (Aditiya HB *et al.*, 2016). Lignocellulosic agricultural and forestry waste materials contain 20–40% of hemicellulose, which is a branched heteropolymer consisting of pentose (D-xylose and L-arabinose) and hexose (D-mannose, D-glucose, and D-galactose) sugars with xylose being most abundant (Bhalla A *et al.*, 2015). Among the different energy-crops for lignocellulosic materials, one of the most interesting is giant reed (*Arundo donax*), because it is tolerant to a wide range of environmental stresses so that it can be cultivated on marginal, degraded or contaminated lands thus reducing competition with food crops which generally require a better quality arable land (Di Nasso NNO *et al.* 2013; Fiorentino N *et al.* 2010; Diodato N *et al.* 2009). *A. donax* hemicellulose is composed by 4-*O*-methyl-glucuronoarabinoxylan and its efficient enzymatic hydrolysis requires a complete repertoire of biomass deconstruction enzymes. Among these, endo- β -xylanases hydrolyze the xylan backbone, α -glucuronidases debranch the side-chain of 4-*O*-methyl-glucuronide and α -arabinofuranosidase remove the arabinoside ramification. In addition, acetyl xylan esterases and feruloyl-esterase are required for a complete hydrolysis. One part of my PhD project was aimed to identify novel thermostable and efficient hemicellulases from thermophilic bacteria in order to achieve a robust enzymatic preparation for the

conversion of glucuronoxylan, the hemicellulose present in *A. donax*.

Xylanases (EC 3.2.1.8 and .32) are currently classified in families GH5, GH8, GH10, GH11, GH26, GH30, GH43 and GH51. Among these, GH10 enzymes, whose catalytic apparatus and three-dimensional structures are well known, cleave glucuronoxylan chains when MeGlcA is linked to xylose at the +1 subsites, (Lombard V *et al.*, 2014; Collins T *et al.*, 2005; Kolenova K *et al.*, 2006). GH67 α -glucuronidases remove glucuronic acid from the non-reducing end of xyloligosaccharides, corresponding to hydrolysis products of GH10. Therefore the xylanases from GH10 are the ideal candidates for hydrolysis of glucuronoxylan in cooperation with α -glucuronidases from GH67 family. Following these considerations, we focused our search among thermophilic enzymes in GH10 and GH67. We selected two ORFs from *Alicyclobacillus acidocaldarius* codifying for a putative endo- β -xylanase (Aaci_2328) and a putative α -glucuronidase (Aaci_0060). Moreover, we identified the ORF Csac_2689 from *Caldicellulosiruptor saccharolyticus* codifying for a putative α -glucuronidase. We performed a multialignment of aminoacid sequences of Aaci_2328, Aaci_0060 and Csac_2689 with all characterized members of respective families. Aaci_2328 showed a low percent of identity (62%) to a characterized GH10 xylanase (Table 2.4.1), while Aaci_0060 and Csac_2689 showed up to 57% and 77% amino acid sequence identity to characterized GH67, respectively (Tables 2.4.2 and .3). These low identities might reflect different structural characteristics and substrate specificity, thus, we embarked on their biochemical characterization.

2.4- Results and discussion

GH10 Characterized	% of identity
gi 263199985 gb ACY69980.1 xylanase [Alicyclobacillus sp. A4]	61.83
gi 444189318 gb AGD81833.1 endo-beta-1,4-xylanase [Geobacillus stearothermophilus]	53.17
gi 291500843 gb ADE08352.1 intracellular GH10 xylanase [Cohnella laeviribosi]	52.37
gi 73332107 gb AAZ74783.1 xylanase [Geobacillus stearothermophilus]	51.66
gi 61287936 dbj BAA31551.2 exo-beta-1,4-xylanase [Aeromonas caviae]	51.35
gi 499714 dbj BAA05668.1 endo-1,4-beta-xylanase [Geobacillus stearothermophilus]	50.00
gi 544636931 gb EPR27574.2 Endo-1,4-beta-xylanase [Geobacillus sp. WSUCF1]	50.00
gi 313574194 dbj BAJ41040.1 xylanase [Paenibacillus curdlanolyticus]	49.41
gi 263199294 gb ACY69972.1 endoxylanase [Paenibacillus sp. E18]	49.24
gi 345548845 gb AE012683.1 xylanase [Paenibacillus xylanilyticus]	49.11
gi 263199944 gb ACY69979.1 xylanase [Anoxybacillus sp. E2(2009)]	49.09
gi 380837255 gb AFE82288.1 endo-1,4-beta-xylanase [Bacillus sp. HJ2]	48.94
gi 2980618 emb CAA76420.1 endo-1,4-beta-xylanase [Thermobacillus xylanilyticus]	48.82
gi 210076633 gb ACJ06666.1 xylanase, partial [Paenibacillus sp. HPL-001]	48.34
gi 3201483 emb CAA07174.1 endo-1,4-beta-xylanase [Paenibacillus barcinonensis]	48.19
gi 429843869 gb AGA16736.1 xylanase [Bacillus sp. SN5]	47.34
gi 157862724 gb ABV90486.1 XynA2 [Paenibacillus sp. JDR-2]	46.15
gi 217337127 gb ACK42920.1 Endo-1,4-beta-xylanase [Dictyoglomus turgidum DSM 6724]	45.51
gi 18476191 gb AAL06078.1 beta-1,4-xylanase [uncultured bacterium]	44.67
gi 662884 emb CAA84631.1 endo-beta-1,4-xylanase [Bacillus sp.]	44.41
gi 973983 gb AAA96979.1 beta-1,4-xylanase [Dictyoglomus thermophilum]	44.08
gi 37694736 gb AAQ99279.1 beta-1,4-xylanase [Bacillus alcalophilus]	43.20
gi 145410981 gb ABP67985.1 Endo-1,4-beta-xylanase [Caldicellulosiruptor saccharolyticus DSM 8903]	43.03
gi 144296 gb AAA23059.1 xylanase/beta-xylosidase (XynA) precursor [Caldicellulosiruptor saccharolyticus]	42.90
gi 224995896 gb ACN76857.1 family 10 endo-beta-xylanase [Glaciecola mesophila KMM 241]	42.90
gi 222455075 gb ACM59337.1 Endo-1,4-beta-xylanase [Caldicellulosiruptor bescii DSM 6725]	42.43
gi 57639627 gb AAW55667.1 xylanase [uncultured organism]	41.72
gi 190686602 gb ACE84280.1 endo-1,4-beta-xylanase, xyn10D [Cellvibrio japonicus Ueda107]	41.72
gi 335365741 gb AEH51686.1 Endo-1,4-beta-xylanase [Thermotoga thermarum DSM 5069]	41.42
gi 21110686 gb AAM39084.1 endo-1,4-beta-xylanase A [Xanthomonas axonopodis pv. citri str. 306]	40.53
gi 307543323 gb ADN44274.1 xylanase [uncultured bacterium]	40.24

Table 2.4.1: % of identity of Aaci_2328 against all GH10 characterized. Here, only the sequences that share more than 40% of identity with Aaci_2328 was reported.

2.4- Results and discussion

GH67 characterized	% of identity vs Aaci_0060
gi 387763702 gb AFJ94648.1 alpha-glucuronidase [uncultured bacterium]	57.25
gi 16930794 gb AAL32057.1 AF441188_1 alpha-glucuronidase [Geobacillus stearothermophilus]	55.08
gi 114054548 gb ABI49940.1 alpha-glucuronidase [Geobacillus stearothermophilus]	54.93
gi 9963829 gb AAG09715.1 AF221859_1 alpha-glucuronidase [Geobacillus stearothermophilus]	54.06
gi 157862723 gb ABV90485.1 AguA [Paenibacillus sp. JDR-2]	53.57
gi 4204209 dbj BAA74508.1 alpha-glucuronidase [Aeromonas caviae]	51.16
gi 343964732 gb AEM73879.1 Glycosyl hydrolase 67 middle domain protein [Caldicellulosiruptor lactoaceticus 6A]	50.58
gi 222455707 gb ACM59969.1 Alpha-glucuronidase [Caldicellulosiruptor bescii DSM 6725]	50.43
gi 217337193 gb ACK42986.1 Alpha-glucuronidase [Dictyoglomus turgidum DSM 6724]	50.22
gi 391417910 gb AFM44650.1 Agu67A [Caldanaerobius polysaccharolyticus]	48.99
gi 332377076 gb AEE64776.1 Agu67A [Ruminococcus albus 8]	46.77
gi 347954065 gb AEP33615.1 alpha-glucuronidase 1 [Chrysosporium lucknowense]	43.33
gi 74592215 sp Q5AQZ4.1 AGUA_EMENI RecName: Full=Alpha-glucuronidase A; AltName: Full=Alpha-glucosiduronase A; Flags: Precursor	42.61
gi 14018210 emb CAC38119.1 alpha-glucuronidase [Aspergillus niger]	39.71
gi 40557182 gb AAR87862.1 alpha glucuronidase [Aureobasidium pullulans]	39.71
gi 134081886 emb CAK42141.1 alpha-glucuronidase aguA-Aspergillus niger	39.71
gi 242332623 emb CAZ66755.1 unnamed protein product [Penicillium aurantiogriseum]	39.71
gi 1419338 emb CAA92949.1 alpha-glucuronidase [Trichoderma reesei]	39.13
gi 32172670 gb AAL33576.3 alpha-glucuronidase [Rasamsonia emersonii]	38.12
gi 190685790 gb ACE83468.1 alpha-glucuronidase, gla67A [Cellvibrio japonicus Ueda107]	37.68
gi 380749163 gb AFE48530.1 alpha-glucuronidase [uncultured rumen bacterium]	37.11
gi 298239738 gb ADI70674.1 alpha-glucuronidase [Prevotella bryantii B14]	37.04
gi 18086518 gb AAL57753.1 alpha-glucuronidase [Cellvibrio mixtus]	34.78

Table 2.4.2: % of identity of Aaci_0060 against all GH67 characterized

GH67 characterized	% of identity vs Csac_2689
gi 222455707 gb ACM59969.1 Alpha-glucuronidase [Caldicellulosiruptor bescii DSM 6725]	77.49
gi 343964732 gb AEM73879.1 Glycosyl hydrolase 67 middle domain protein [Caldicellulosiruptor lactoaceticus 6A]	77.34
gi 217337193 gb ACK42986.1 Alpha-glucuronidase [Dictyoglomus turgidum DSM 6724]	60.73
gi 391417910 gb AFM44650.1 Agu67A [Caldanaerobius polysaccharolyticus]	58.84
gi 16930794 gb AAL32057.1 AF441188_1 alpha-glucuronidase [Geobacillus stearothermophilus]	57.44
gi 114054548 gb ABI49940.1 alpha-glucuronidase [Geobacillus stearothermophilus]	57.29
gi 387763702 gb AFJ94648.1 alpha-glucuronidase [uncultured bacterium]	55.46
gi 157862723 gb ABV90485.1 AguA [Paenibacillus sp. JDR-2]	51.09
gi 9963829 gb AAG09715.1 AF221859_1 alpha-glucuronidase [Geobacillus stearothermophilus]	50.22
gi 332377076 gb AEE64776.1 Agu67A [Ruminococcus albus 8]	48.31
gi 4204209 dbj BAA74508.1 alpha-glucuronidase [Aeromonas caviae]	47.70
gi 347954065 gb AEP33615.1 alpha-glucuronidase 1 [Chrysosporium lucknowense]	38.94
gi 74592215 sp Q5AQZ4.1 AGUA_EMENI RecName: Full=Alpha-glucuronidase A; AltName: Full=Alpha-glucosiduronase A; Flags: Precursor	38.79
gi 242332623 emb CAZ66755.1 unnamed protein product [Penicillium aurantiogriseum]	38.22
gi 1419338 emb CAA92949.1 alpha-glucuronidase [Trichoderma reesei]	37.93
gi 134081886 emb CAK42141.1 alpha-glucuronidase aguA-Aspergillus niger	37.93
gi 14018210 emb CAC38119.1 alpha-glucuronidase [Aspergillus niger]	37.79
gi 40557182 gb AAR87862.1 alpha glucuronidase [Aureobasidium pullulans]	37.50
gi 298239738 gb ADI70674.1 alpha-glucuronidase [Prevotella bryantii B14]	36.41
gi 32172670 gb AAL33576.3 alpha-glucuronidase [Rasamsonia emersonii]	36.21
gi 190685790 gb ACE83468.1 alpha-glucuronidase, gla67A [Cellvibrio japonicus Ueda107]	35.92
gi 18086518 gb AAL57753.1 alpha-glucuronidase [Cellvibrio mixtus]	35.49
gi 380749163 gb AFE48530.1 alpha-glucuronidase [uncultured rumen bacterium]	34.01

Table 2.4.3: % of identity of Csac_2689 against all GH67 characterized

The presence of a possible signal peptides at the amino-terminal of each putative enzyme was searched by using the prediction programs PRED-TAT and SignalP 4.0. Both programs did not reveal any secretion peptide and proteolytic signalling sequence in any of the polypeptides analyzed, suggesting that these proteins were intracellular.

Cloning, expression and purification

To determine their biochemical characteristics and their catalytic activities, we produced and characterized the enzymes in recombinant form. To this aim, we cloned the genes Aaci_2328, Aaci_0060 and Csac_2689 in the expression vector pET101/D-TOPO obtaining the recombinant constructs pET101/D-TOPO-Aaci_2328, pET101/D-TOPO- Aaci_0060 and pET101/D-TOPO-Csac_2689, which allowed to clone the genes in frame with a 6His-tag at the carboxy-terminal of the enzymes of interest, greatly facilitating their purification by affinity chromatography. The recombinant clones were sequenced showing some differences compared to that reported in database. In particular, the sequence of Aaci_2328 differed in 13 nucleotides; however, only one of these changes determined an aminoacidic substitution (Gly512Asp), which has been confirmed in two independent clones and is conserved in homolog enzymes from different *A. acidocaldarius* strains (Figure 2.4.1). Instead, the sequence of Aaci_0060 differed from the one deposited in GenBank at the NCBI in about 100 nucleotides, determining 34 aminoacidic substitutions, 13 of which conserved in the *A. acidocaldarius* Tc4-1 strain (Figure 2.4.2). The DNA sequences of both these ORFs were deposited in GenBank at the NCBI with the accession numbers KJ466334 and KJ466335, respectively. The positive clones of Csac_2689 did not show mutations compared to sequence in databank.

2.4- Results and discussion

```
Aaci_2328      ATGACGGATCAAGCGCGCTCTCTGAAAGAGCGTACGCTTCCCGCTTTTCGCGTGGCGCT 60
Aaci_2328_DB   ATGACGGATCAAGCGCGCTCTCTGAAAGAGCGTACGCTTCCCGCTTTTCGCGTGGCGCT 60
Strain_Tc-4-1  ATGACGGATCAAGCGCGCTCTCTGAAAGAGCGTACGCTTCCCGCTTTTCGCGTGGCGCT 60
Strain_A15     ATGACGGATCGAGTCCATCTGTCGCGAAGCGTATGCGCGCATTTTCGCGTGGCGCGC 60
*****,**.**,** *...***** ** *..***** *****

Aaci_2328      GCGGTCAACGCGCGGACCGTTTACACGACGACGCCATCTCTGGCGCGCCACTTCAGCAGT 120
Aaci_2328_DB   GCGGTCAACGCGCGGACCGTTTACACGACGACGCCATCTCTGGCGCGCCACTTCAGCAGT 120
Strain_Tc-4-1  GCGGTCAACGCGCGGACCGTTTACACGACGACGCCATCTCTGGCGCGCCACTTCAGCAGT 120
Strain_A15     GCGGTCAACGCGCGGACCGTTTACACGACGACGCCATCTCTGGAGGTCACTTCAGCAGT 120
***** ** * ..***** ***** ** * ..***** *****

Aaci_2328      GTGACCGCGGAGAACGAGATGAAGTGGGAGCGCATCCACCCGCGGAGAACACGTATTCT 180
Aaci_2328_DB   GTGACCGCGGAGAACGAGATGAATGGGAGCGCATCCACCCGCGGAGAACACGTATTCT 180
Strain_Tc-4-1  GTGACCGCGGAGAACGAGATGAATGGGAGCGCATCCACCCGCGGAGAACACGTATTCT 180
Strain_A15     ATCACCGCGGAGAACGAGATGAAGTGGGAGCGCATCCACCCGCGGAGAACACGTATTCT 180
*, ***** ***** ** * ..**,* *****

Aaci_2328      TTTTCTGCTGCTGACCAATCGTTTATTTCGCCCGGATCAGGCGATGTTCTGCGCGCG 240
Aaci_2328_DB   TTTTCTGCTGCTGACCAATCGTTTATTTCGCCCGGATCAGGCGATGTTCTGCGCGCG 240
Strain_Tc-4-1  TTTTCTGCTGCTGACCAATCGTTTATTTCGCCCGGATCAGGCGATGTTCTGCGCGCG 240
Strain_A15     TTTTCTGCTGCTGACCAATCGTTTATTTCGCCCGGATCAGGCGATGTTCTGCGCGCG 240
***** ** * ..***** ***** ** * ..***** *****

Aaci_2328      CACACGCTCGTGTGGCACAACAGACGCCCTTCTGGGTGTTCTTGAGCTCTCTCGGCCAA 300
Aaci_2328_DB   CACACGCTCGTGTGGCACAACAGACGCCCTTCTGGGTGTTCTTGAGCTCTCTCGGCCAA 300
Strain_Tc-4-1  CACACGCTCGTGTGGCACAACAGACGCCCTTCTGGGTGTTCTTGAGCTCTCTCGGCCAA 300
Strain_A15     CACACGCTCGTGTGGCACAATCAGACGCCCGGTGGGTGTTCTCGGCGAGCGCTCGGCCAA 300
***** ** * ..***** ***** ** * ..***** *****

Aaci_2328      CCGCGCGCGCGAAGCTCGTGGAGCGAAGGCTCGAGCAGCATCGCGGAGGTCTGGGGT 360
Aaci_2328_DB   CCGCGCGCGCGAAGCTCGTGGAGCGAAGGCTCGAGCAGCATCGCGGAGGTCTGGGGT 360
Strain_Tc-4-1  CCGCGCGCGCGAAGCTCGTGGAGCGAAGGCTCGAGGCGCATCGCGGAGGTCTGGGGG 360
Strain_A15     TCGCGCGCGCGAAGCTCGTGGAGGAGGCTCGAGGCGCATCGTCTGAGGTCTGCGGG 360
* ***** ** * ..***** ***** ** * ..***** *****

Aaci_2328      CACTACCGCGCGCGCGCTTGTGCTGGGACGTGTGAACGAGGCGGTGATCGACCAAGGG 420
Aaci_2328_DB   CACTACCGCGCGCGCGCTTGTGCTGGGACGTGTGAACGAGGCGGTGATCGACCAAGGG 420
Strain_Tc-4-1  CACTACCGCGCGCGCGCTTGTGCTGGGACGTGTGAACGAGGCGGTGATCGACCAAGGG 420
Strain_A15     CACTACCGCGCGCATGTGCGGTGCTGGGACGTGTGAACGAGGCGGTGATCGACCAAGGG 420
***** ** * ..***** ***** ** * ..***** *****

Aaci_2328      GACGCGTGGCTGCGCGCGAGCCATGCGCGCAGGCGCTGGGGACGACTACATCGAGATG 480
Aaci_2328_DB   GACGCGTGGCTGCGCGCGAGCCATGCGCGCAGGCGCTGGGGACGACTACATCGAGATG 480
Strain_Tc-4-1  GACGCGTGGCTGCGCGCGAGCCATGCGCGCAGGCGCTGGGGACGACTACATCGAGAG 480
Strain_A15     GACGCGTGGCTTCCGCCAAGCCATGCGCGCAGGCGCTCGGGACGATTACATCGAGAGA 480
***** ** * ..***** ***** ** * ..***** *****

Aaci_2328      GCGTTTCGCTTGGCGCACCAAGCGGATCCCGACGCTCTCTTCTACAACGACTAAAC 540
Aaci_2328_DB   GCGTTTCGCTTGGCGCACCAAGCGGATCCCGACGCTCTCTTCTACAACGACTAAAC 540
Strain_Tc-4-1  GCGTTTCGCTTGGCGCACCAAGCGGATCCCGACGCTCTCTTCTACAACGACTAAAC 540
Strain_A15     GCGTTTCGCTGCGCGCACCAAGCGGATCCCGACGCTCTCTTCTACAACGACTAAAGT 540
***** ** * ..***** ***** ** * ..***** *****

Aaci_2328      GAGACGAAGCCGAGAAAGCGGAGCCGATCTCTCGCGGTGCTCGAGCACCTGTTGGACCG 600
Aaci_2328_DB   GAGACGAAGCCGAGAAAGCGGAGCCGATCTCTCGCGGTGCTCGAGCACCTGTTGGACCG 600
Strain_Tc-4-1  GAGACGAAGCCGAGAAAGCGGAGCCGATCTCTCGCGGTGCTCGAGCACCTGTTGGACCG 600
Strain_A15     GAGACGAAGCCTTTAAGCGAGACCGCATCTCTCGCGGTGCTAGAGCACCTGTTGGACCG 600
***** ** * ..***** ***** ** * ..***** *****

Aaci_2328      GGAAGTCCCGGTGATGCGGTGGGGCTGAGATGACAGTCTCCCTGGAGCATCCGCCATC 660
Aaci_2328_DB   GGAAGTCCCGGTGATGCGGTGGGGCTGAGATGACAGTCTCCCTGGAGCATCCGCCATC 660
Strain_Tc-4-1  GGAAGTCCCGGTGATGCGGTGGGGCTGAGATGACAGTCTCCCTGGAGCATCCGCCATC 660
Strain_A15     GCGGTGCCATTCATGCGGTGGGGCTGAGATGACAGTCTCCCTGGAGCATCCGCCATC 660
**,* ** * ..***** ***** ** * ..***** *****
```

Figure 2.4.1: Alignment between Aaci_2328 sequenced in this study (Aaci_2328), Aaci_2328 present in database (Aaci_2328_DB), GH10 from *A. acidocaldarius* Tc-4-1 (Strain_Tc-4-1) and GH10 from *A. acidocaldarius* A15 (Strain_A15). In yellow the mutations of Aaci_2328 of this study, in green the mutation of the sequence in database. The arrows indicate the conserved mutations. The red box highlights the mutation determining aminoacidic substitution.

2.4- Results and discussion



Figure 2.4.2: Alignment between Aaci_0060 sequenced in this study (Aaci_0060), Aaci_0060 present in database (Aaci_0060_DB) and GH67 from *A. acidocaldarius* Tc-4-1 (Strain_Tc-4-1). The mutations of Aaci_0060 identified in this study and those of the sequence in database are highlighted in yellow and green, respectively. Arrows indicate the conserved mutations. In red box the mutations that determined the aminoacidic substitutions.

To determine which conditions allowed to obtain the best yields of each enzyme, expression trials were performed in *E. coli* BL21 star (DE3). SDS-PAGE analysis showed that the best expression levels for Aaci_2328 could be obtained after induction with 0.1 mM IPTG for 16 hours when the growth reached 1 OD. While the best yields of Aaci_0060 and Csa_2689 could be achieved inducing for 16 hours with 0.5 mM IPTG when the growth reached 1 OD. The purification procedures consisted in an IMAC chromatography step, followed by incubation at high temperature and removal of aggregated proteins

and, in the case of GH67s, and a S-200 gel filtration chromatography. This purification procedure led to a purity grade of 95%, 80% and 99% for Aaci_2328, Aaci_0060 and Csac_2689, respectively, evaluated by SDS-PAGE (Figure 2.3.3). The recombinant enzymes were named GH10-XA, GH67-GA and GH67-GC.

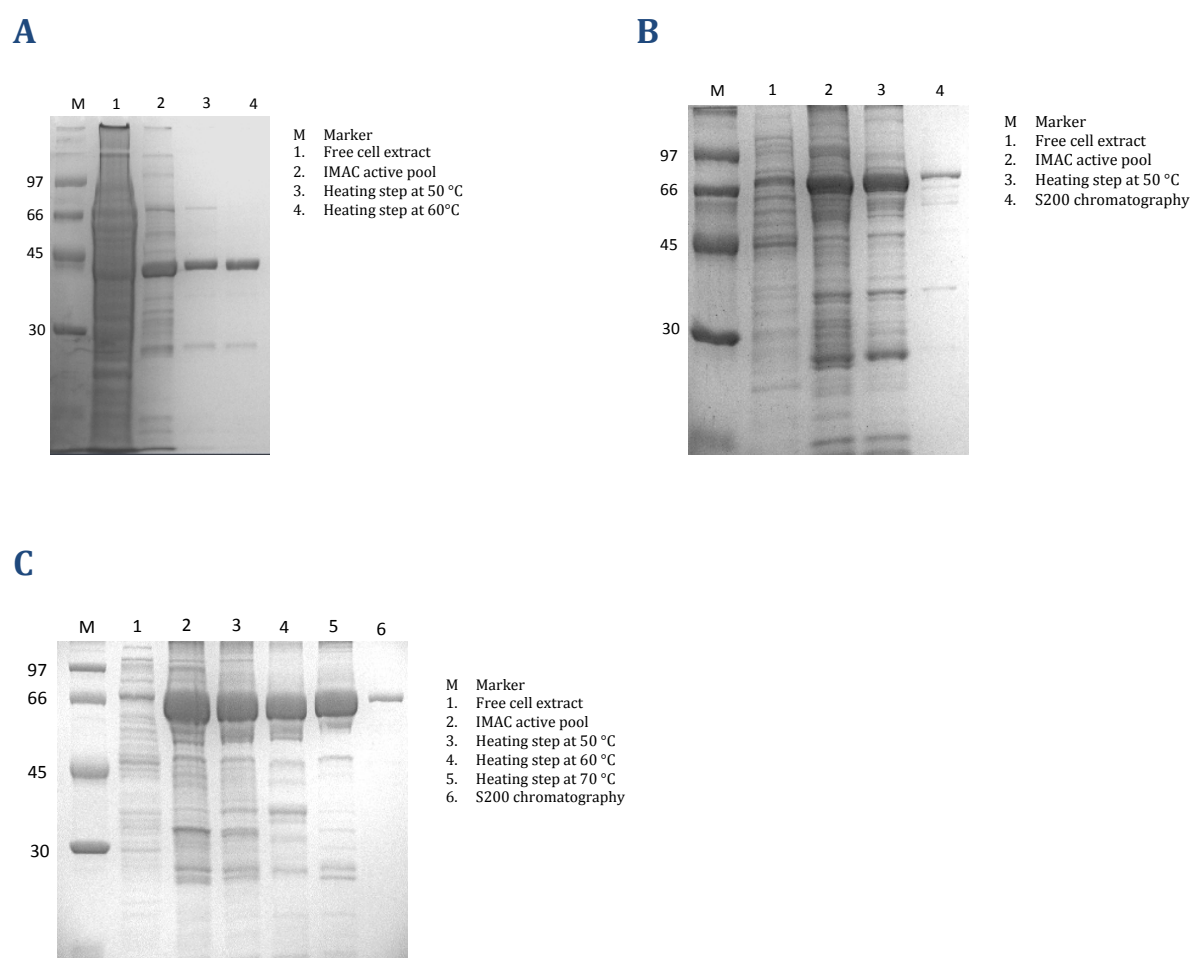


Figure 2.4.3: SDS-PAGE of the purification steps of GH10-XA (A), GH67-GA (B) and GH67-GC (C).

Both GH10-XA and GH67-GC displayed good final yields corresponding to 6 mg and 3.1 mg per liter of *E. coli*, respectively, whereas GH67-GA resulted in only 0.6 mg per liter of culture with lower purification yields (Table 2.4.3).

2.4- Results and discussion

A

Purification of GH10 -XA								
Sample	Vol (mL)	Protein concentration (mg mL ⁻¹)	Total proteins (mg)	Enzymatic units (U mL ⁻¹)	Total Enzymatic units (U)	Specific activity (U mg ⁻¹)	Yields (%)	Purification fold
Free cell extract	44.5	15.2	677	106.5	4739	7	100	1
His Trap	6.0	4.2	25	360.0	2150	86	45	12
Heating step I (50°C)	5.6	3.7	21	364.0	2037	97	43	14
Heating step II (60°C)	5.4	2.2	12	233.0	1260	105	27	15

B

Purification of GH67 -GA								
Sample	Vol (mL)	Protein concentration (mg mL ⁻¹)	Total proteins (mg)	Enzymatic units (U mL ⁻¹)	Total Enzymatic units (U)	Specific activity (U mg ⁻¹)	Yields (%)	Purification fold
Free cell extract	35	17.4	610	1.6	54.9	0.1	100	1
His Trap	4	8.1	32.2	6.7	26.7	0.8	49	9
Heating step (50°C)	2	2.3	4.6	5.7	11.5	2.5	21	28
S-200 chromatography	2.5	0.50	1.2	2.1	5.3	4.39	10	49

C

Purification of GH67 -GC								
Sample	Vol (mL)	Protein concentration (mg mL ⁻¹)	Total proteins (mg)	Enzymatic units (U mL ⁻¹)	Total Enzymatic units (U)	Specific activity (U mg ⁻¹)	Yields (%)	Purification fold
Free cell extract	34	18.98	645	1.33	45.17	0.07	100	1
His Trap	5.5	6.54	35.9	4.70	25.85	0.72	57	10
Heating step I (50°C)	4.5	5.80	26.1	5.22	23.49	0.90	52	13
Heating step II (60°C)	4.2	5.11	21.5	5.22	21.92	1.02	48	15
Heating step III (70°C)	4	2.98	11.9	4.28	17.14	1.44	38	20
S-200 chromatography	12	0.50	6.2	1.13	13.64	2.21	30	32

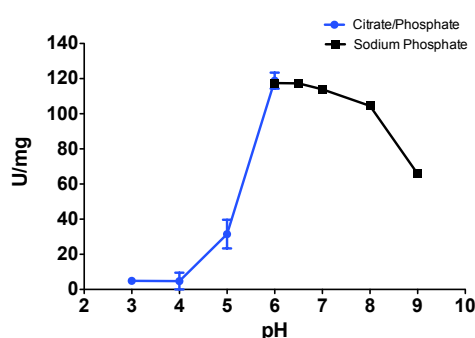
Table 2.4.3: (A) Purification of GH10-XA from *E. coli* BL21(DE3) star with pET101/D-TOPO-Aaci_2328. Assays were performed on 5 mg mL⁻¹ beechwood xylan in 50 mM sodium phosphate buffer pH 6.5 at 65°C for 1 minute. (B) Purification of GH67-GA from *E. coli* BL21(DE3) star with pET101/D-TOPO-Aaci_0060. Assays were performed on 2 mg mL⁻¹ aldouronic acids in 100 mM sodium phosphate buffer pH 6.5 at 50°C for 1 minute. (C) Purification of GH67-GC from *E. coli* BL21(DE3) star with pET101/D-TOPO-Csac_2689. Assays were performed on 1 mg mL⁻¹ aldouronic acids in 100 mM sodium phosphate buffer pH 6.0 at 65°C for 2 minutes.

Biochemical characterization of recombinant enzymes

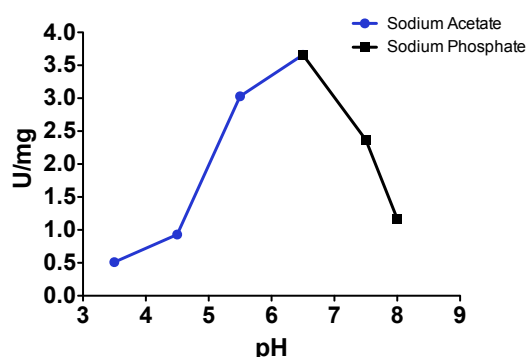
The characterization of GH10-XA, GH67-GA and GH67-GC was performed to unveil the optimal conditions for catalysis, such as pH and temperature optima and thermal stability. To determine the pH optimum of GH10-XA, the enzyme was assayed in 50 mM of different buffers in the range of pH 3.0-9.0 on 5 mg mL⁻¹

of beechwood xylan at 65°C. The hydrolysis products were quantified through the reducing-end assay Somogyi-Nelson, as described in Experimental Procedures. As illustrated in Figure 2.4.4A, GH10-XA was optimally active in sodium phosphate buffer in the pH range 6.0-7.0. The activity at pH near the neutrality indicates that GH10-XA is an intracellular enzyme as *A. acidocaldarius* grows optimally at pH 3.0–4.0. The pH dependence of GH67s was analyzed in the range of pH 3.0-8.0: GH67-GA was assayed on 2 mg ml⁻¹ of aldouronic acids at 50 °C while GH67-GC on 1 mg ml⁻¹ of aldouronic acids at 65 °C. The optimal pHs of the two enzymes were similar (in 100 mM sodium phosphate buffer, pH 6.0 and 6.5 for GH67-GA and GH67-GC, respectively) (Figure 2.4.4B and C), confirming their putative intracellular localization.

A



B



C

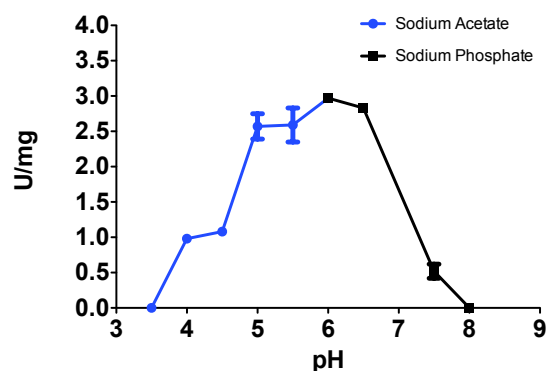


Figure 2.4.4: pH dependence of GH10-XA (A), GH67-GA (B), GH67-GC (C).

As expected for enzymes from thermophilic source, the temperature profile assessed in the range of 40-85°C for GH10-XA (Figure 2.4.5A), 40-65°C for GH67-GA (Figure 2.4.5B), and 40-70°C for GH67-GC (Figure 2.4.5C). The xylanase and the α -glucuronidase from *A. acidocaldarius* were optimally active at 75 °C and at 50°C, respectively while the GH67 from *C. saccharolyticus* showed a maximal activity at 65°C.

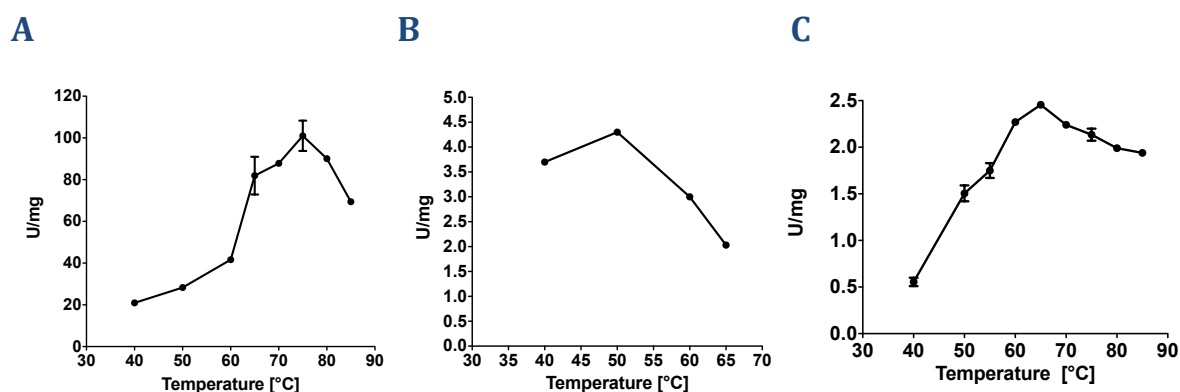
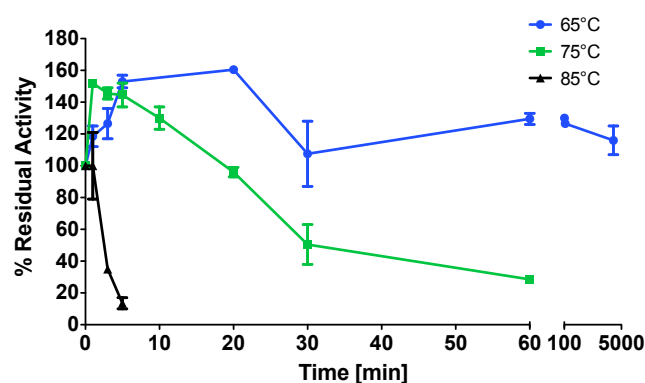


Figure 2.4.5: Dependence from temperature of GH10-XA (A), GH67-GA (B) and GH67-GC (C).

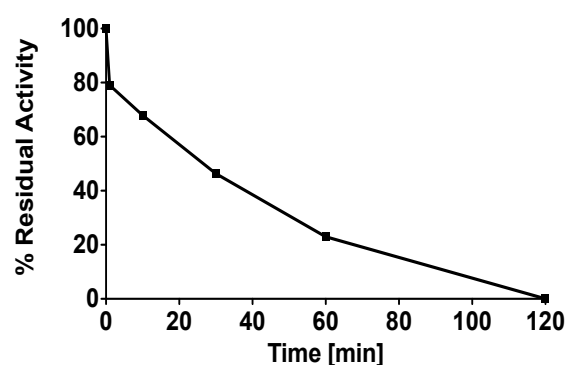
An important feature that makes the thermophilic enzymes applicable for industrial processes is their thermostability. Therefore, we evaluated the stability to heat of each enzyme. GH10-XA incubated at 65°C, 75°C, and 85°C temperatures showed a remarkable stability being still fully active after 3.5 days of incubation at 65 °C and maintaining 40% specific activity after 1 h at 75 °C (Figure 2.3.6A). It is worth noting the remarkable activation observed after 10 min at 65 °C, a phenomenon observed also in other thermophilic enzymes (Cobucci-Ponzano B *et al.*, 2003). The remarkable stability of GH10-XA makes this enzyme a promising tool for several potential applications outcompeting other xylanases. In fact, the GH10 xylanase from a different *A. acidocaldarius* strain, with the highest sequence identity (62%) to GH10-XA, was maximally active at 55 °C and was stable only 20 min at 65 °C (Bai Y *et al.*, 2010) (Table 2.4.5). The specific activity of this enzyme on beechwood xylan was not

reported, debaring a direct comparison with GH10-XA that, on this substrate, is more active than the enzymes from *Thermotoga thermarum*, *Volvariella volvacea*, *Thermoanaerobacterium saccharolyticum*, and *C. lactoaceticus*. In addition, GH10-XA is more stable than the enzymes from *Streptomyces olivaceoviridis*, *Aspergillus niger*, *Bispora*, *Streptomyces* and *Actinomadura* spp, which show higher specific activity on beechwood xylan (Table 2.4.5).

A



B



C

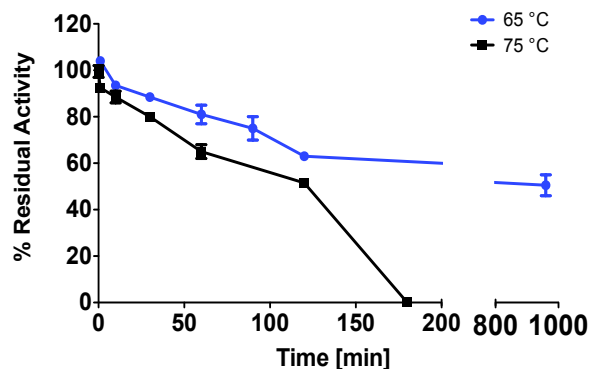


Figure 2.4.6: Thermostability of GH10-XA (A), GH67-GA (B) and GH67-GC (C).

The thermostability of GH67-GA was measured only at 50 °C while GH67-GC was analyzed at 65 °C and 75 °C. GH67-GA resulted unstable to heat by losing 100% of its activity after 120 min at 50 °C (Figure 2.4.6B). By contrast, GH67-GC was remarkably stable: the maintenance of 50% activity after 16 h at 65 °C and 2 h

2.4- Results and discussion

at 75°C makes it one of the most stable thermophilic α -glucuronidases known (Table 2.4.6).

Microorganism	Specific activity	Substrate	Reference	Thermostability	Identity vs GH10-XA
<i>Actinomodura sp. strain Cpt20</i>	712.0 \pm 58.0	beechwood xylan	Tabti, Zina. 2012. Applied Biochemistry and Biotechnology 166 (3) 663–79.	100%(60°C/10h)/ - 30%(70°C/10h)/- 0%(80°C/9h)	Native
<i>A. acidocaldarius (GH10-XA)</i>	152.0 \pm 5.1	beechwood xylan	this study	>100%(65°C/3.5 days) - 50%(75°C/1h) - 10%(85°C/5')	100%
<i>A. acidocaldarius sp. A4</i>	420.2	birchwood xylan	Bai, Yingguo. 2010. Journal of Industrial Microbiology and Biotechnology 37: 187–194.	90%(60°C/1h) - 40%(65°C/1h) – 0%(70°C/10')	62%
<i>A. niger</i>	3,2	beechwood xylan	Zheng, Jia. 2013. Biotechnology letters 1433–1440.	90% (60°C/30') - 15% (70°C/10') – 5% (80°C/5')	21%
<i>Bispora sp MEY-1</i>	2,463	beechwood xylan	Luo, Huiying. 2010. Applied Microbiology and Biotechnology 86 (6) (May): 1829–39.	90%(60°C/1h) - 25%(70°C/10')	18%
<i>C. lactoaceticus</i>	44.6	beechwood xylan	Jia, Xiaojing2014. PloS One 9 (9) (January): e106482. doi:10.1371/journal.pone.0106482.	100%(75°C/6h) -70%(80°C/6h) - 25%(85°C/30')	16%
<i>Marasmius sp</i>	336.0 \pm 22.0	beechwood xylan	Ratanachomr, Ukrit. 2006. Journal of Biochemistry and Molecular Biology 39: 105–110.	95%(50°C/3h) - 80%(60°C/3h) - 54%(70°C/3h)	Native, putative GH5 by taxa
<i>P. thermophila</i>	936	beechwood xylan	Li, Lite 2006. "Enzyme and Microbial Technology 38 (6) (April): 780–787.	95%(70°C/30') - 80%(80°C/30')	Native, putative GH11 by taxa
<i>S. olivaceoviridis E-86</i>	332.5	beechwood xylan	Jiang, Z.Q. 2005Enzyme and Microbial Technology 36 (7) (May): 923–929.	50%(65°C/30')	33%
<i>S. rameus L2001</i>	4326 \pm 97.0	beechwood xylan	Li, Xuting. 2010. Biochemical Engineering Journal 52 (1) (October): 71–78.	90%(65°C/30') - 60%(70°C/30')	21% (GH11)
<i>Streptomyces sp. CS428</i>	926,103.0	beechwood xylan	G.C., Pradeep. 2013. Process Biochemistry 48 (8) (August): 1188–1196.	90%(50°C/1h) - 60%(60°C/1h) - 50%(70°C/1h)	Native
<i>S. racemosum Cohn</i>	1,402.0	birchwood xylan	Sapre, Meenakshi P. 2005. The Journal of General and Applied Microbiology 51 (6): 327–334.	not well characterized	Native
<i>T. saccharolyticum NT0U1</i>	78.0 \pm 4.4	beechwood xylan	Hung, Kuo-Sheng. 2011. Process Biochemistry	65%(60°C/2h) - 30%(65°C/2h) - 10%(75°C/10')	16%
<i>T. lanuginosus CBS 288.54</i>	895.0 \pm 21.6	beechwood xylan	Li, X T 2005. Bioresource Technology 96 (12) (August): 1370–9.	90%(70°C/30') -60%(80°C/30')	Native, putative GH11 by taxa
<i>T. thermarum</i>	145.8	beechwood xylan	Shi, Hao. 2013. Biotechnology for Biofuels 6: 26.	100%(85°C/2h) -30%(90°C/2h) -0%(95°C/60')	41%
<i>V. volvacea</i>	67.3 \pm 0.8	beechwood xylan	Zheng, Fei. 2013. Journal of Industrial Microbiology & Biotechnology 40 (10) (October): 1083–93.	70%(50°C/60') - 10%(55°C/60')	23%

Table 2.4.5: Comparison between GH10-XA and characterized xylanases.

Microorganism	Specific activity	Substrate	Reference	Thermostability	Identity vs GH67-GA	Identity vs GH67-GC
<i>A. acidocaldarius</i>	4.6 \pm 0.3	Aldouronic acids	this study	50% (50°C/30min)	100	49.42
<i>C. saccharolitycus</i>	2.6 \pm 0.2	Aldouronic acids	this study	50% (65°C/16h)- 50% (75°C/2h)	49.42	100
<i>G. stereothermophilus T-6</i>	42.0	Aldouronic acids	Zaide G. Eur. J. Biochem. 268, 3006-3016 (2001) e1835 2001	100% (65°C/20 min) - 100% (70°C/20 min); 30% (75°C/20 min)	54.93	57.29
<i>G. stereothermophilus 236</i>	14.2	Aldouronic acids	Choi I-D. Biosci. Biotechnol. Biochem., 64(12), 2530-2537, 2000	not well characterized	54.06	50.22
<i>Paenibacillus sp. Strain JDR-2:</i>	5.54	Aldouronic acids	Guang Nong. 2009. Applied and environmental microbiology.	not well characterized	53.57	51.09
<i>C. lactoaceticus</i>	1.3	Aldobiuronic acid	Jia, Xiaojing2014. PloS One 9 (9) (January): e106482. doi:10.1371/journal.pone.0106482	not well characterized	50.58	77.34
<i>C. polysaccharolyticus</i>	154	Aldouronic acids	Yejun Han. J Biol Chem. 2012 Oct 12; 287(42): 34948–34960.	not well characterized	48.99	58.84
<i>A. pullulans</i>	14.1 \pm 0.08; 60 \pm 4.3; 135 \pm 6.2; 126 \pm 14	Aldobiouronic acid Aldotriuronic acid Aldotetrauronic acid Aldopentaauronic acid	B.J.M. de Wet. 2006. Enzyme and Microbial Technology 38 (2006) 649–656	50% (55°C/23 min); 50% (65°C/6 min)	39.71	37.50
<i>S. commune</i>	2.3	Aldotriuronic acid	Tenkanen M. 2000. Journal of Biotechnology 78 (2000) 149–161	65% (45°C/24 h)	native	native
<i>P. stiptitis</i>	6.2	Aldopentaauronic acid	Ryabova O. 2009. FEBS Letters 583 (2009) 1457–1462A	stable (40°C/3h)- 50% (60°C 30 min)	native	native
<i>P. radiata</i>	1389.96	Aldo-bi/triouronic acid	Malgorzata Mierzwia . Preparative Biochemistry and Biotechnology.	not well characterized	native	native
<i>C. japonicus</i>	0.058 \pm 0.8 * 10 ⁻⁴ 0.061 \pm 4.1*10 ⁻³ 0.04 \pm 3.0*10 ⁻³ 0.084 \pm 11.7*10 ⁻³	Aldobiouronic acid Aldotriuronic acid Aldotetrauronic acid Aldopentaauronic acid	Tibor Nagy. Journal of Bacteriology, Sept. 2002, .	rapidly inactivated at 55 °C	native	native
<i>Uncultured bacterium</i>	11.7	Aldouronic acids	Zheng, Fei. 2013. Journal of Industrial Microbiology & Biotechnology 40 (10) (October): 1083–93.	75% (40°C/120 min)- 70% (50°C / 120 min)- 0% (60°C/ 60 min)	59.15	57.97

Table 2.4.6: Comparison between GH67-GA and –GC and characterized GH67.

On the basis of this general characterization, the standard assays used in all the following study was summarized in Table 2.4.7.

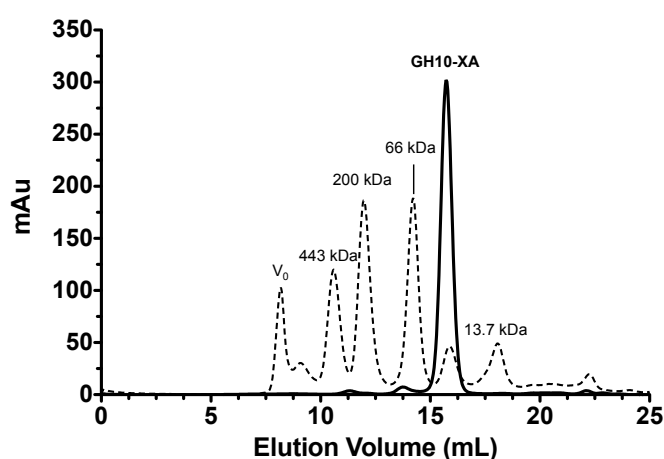
2.4- Results and discussion

Enzyme	Substrate	pH	Temperature (°C)	Time (min)
GH10-XA	5 mg mL ⁻¹ beechwood xylan	50 mM sodium phosphate pH 6.5	65	1
GH67-GA	2 mg mL ⁻¹ Aldouronic acids	100 mM sodium phosphate pH 6.5	50	1
GH67-GC	1 mg mL ⁻¹ Aldouronic acids	100 mM sodium phosphate pH 6.0	65	2

Table 2.4.7: Conditions for standard assays.

The molecular weights of the enzymes were analyzed by size exclusion chromatography in native conditions. The xylanase showed a molecular mass of 39 ± 2 kDa indicating a monomeric structure (Figure 2.4.7A). The gel-filtration of GH67-GA and GH67-GC revealed molecular masses of 153 ± 2 and 109 ± 18 kDa, respectively (Figure 2.4.7B and C). Since the molecular masses in denatured conditions were about 80 kDa, this might indicate that the native enzymes were either relaxed monomers or, by analogy with all the other GH67 enzymes, tight dimers (Ruile C *et al.*, 1997)

A



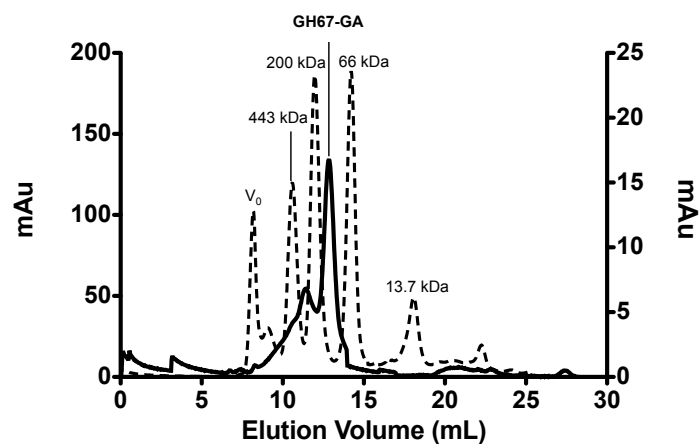
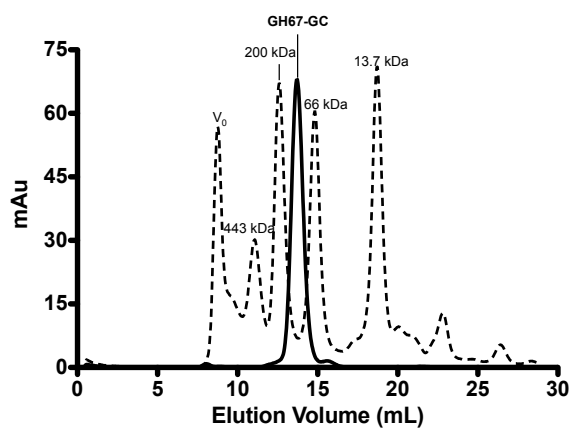
B**C**

Figure 2.4.7: Molecular mass determination of GH10-XA (**A**), GH67-GA (**B**), GH67-GC (**C**).

Substrate specificity

GH10-XA

In order to determine the substrate specificity of GH10-XA, this enzyme was tested on 4-*O*-methylglucuronoxylan (MGX), beechwood xylan, 2Np- β -cellobioside (2Np-Cel), 4Np- β -xylopyranoside (4Np-Xyl), 4Np- β -glucopyranoside (4Np-Glc), carboxy-methyl-cellulose (CMC), Avicel, cellobiose, and cellotriose, as described in Experimental procedures ([Table 2.4.8](#)).

Substrates	Specific activity (U mg ⁻¹)
MGX	170.4 ± 7.6
Beechwood xylan	152.0 ± 5.1
2Np-Cel	86.9 ± 3.2
4Np-Xyl	6.7 ± 0.4
4Np-Glc	ND
CMC	ND
Avicel	ND
cellobiose ¹	ND
cellotriose ¹	ND

Table 2.4.8: Assays of GH10-XA on different substrates. ND: not detectable. ¹determined by TLC

As expected from a GH10 enzyme (Kim DY *et al.*, 2014), GH10-XA showed clear specificity toward xylan substrates (170.4 ± 7.6 U mg⁻¹ on MGX and 152.0 ± 5.1 U mg⁻¹ on beechwood xylan), but also it was considerably active on 2Np-Cel (86.9 ± 3.2 U mg⁻¹) but not on 4Np-Xyl (6.7 ± 0.4 U mg⁻¹) (Table 2.4.8).

The steady state kinetic constant of GH10-XA on MGX, beechwood xylan, 2Np-Cel and 4Np-Xyl were measured in standard conditions and listed in Table 2.4.9.

Substrate	k _{cat} (s ⁻¹)	K _M (mg mL ⁻¹)	k _{cat} /K _M (s ⁻¹ mg ⁻¹ mL)
Beechwood xylan	98.2±4.1	1.81±0.16	54.3
MGX	110.8±5.0	0.79±0.12	140.3
2Np-Cel	56.5±2.1	0.51±0.08	110.8
4Np-Xyl	4.4±0.3	1.62±0.30	2.7

Table 2.4.9: Kinetic constants of GH10-XA on different substrates.

The highest specificity constant was observed on MGX followed by 2Np-Cel and beechwood xylan while the enzyme confirmed its barely activity on 4Np-Xyl (Table 2.4.9). However, it is worth mentioning that GH10-XA incubated in the

presence of high concentrations of this substrate (30 mM) yielded transglycosylation products observable by TLC (Figure 2.4.8). Among GH10 enzymes, only XynB from *Paenibacillus* sp catalyzed transglycosylation reactions, but from different substrates (xylotriose and xylotetraose) (Blanco A *et al.*, 1996). Moreover, for the first time in GH10 family, GH10-XA catalyzed the formation of transglycosylation products also in the presence of aryl-glycoside substrates (4Np-Glc and 2Np-Cel at 5 and 8 mM, respectively) (Figure 2.4.8 and 2.4.9), while cellobiose and cellotriose were not substrates of the enzyme (Table 2.4.8).

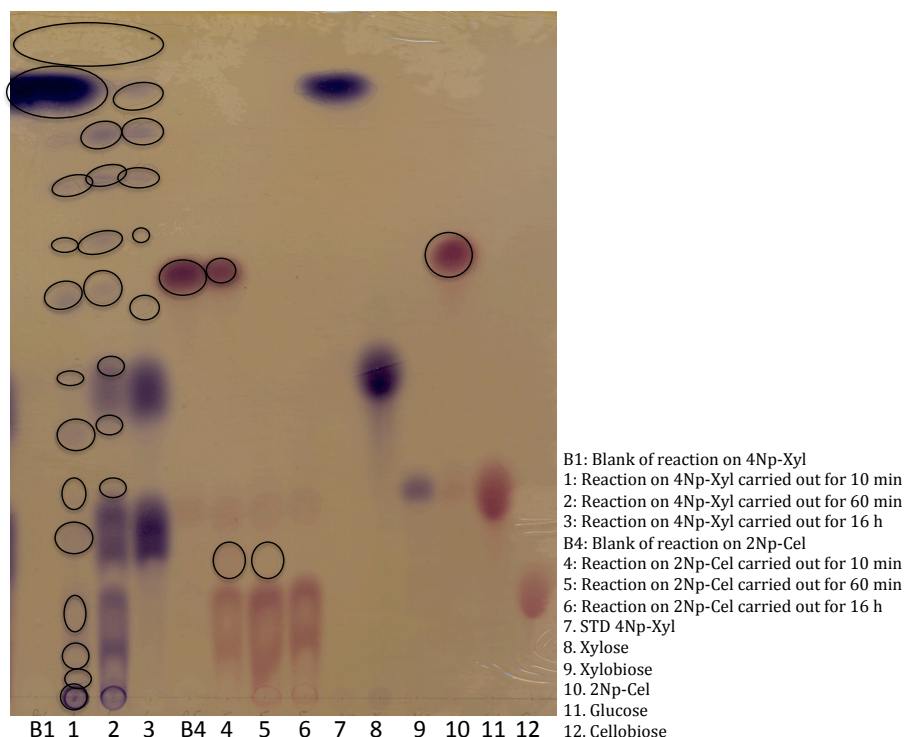


Figure 2.4.8: TLC analysis of enzymatic assays on 4Np-Xyl (30 mM) and 2Np-Cel (8 mM). The circle indicated the UV-visible products.

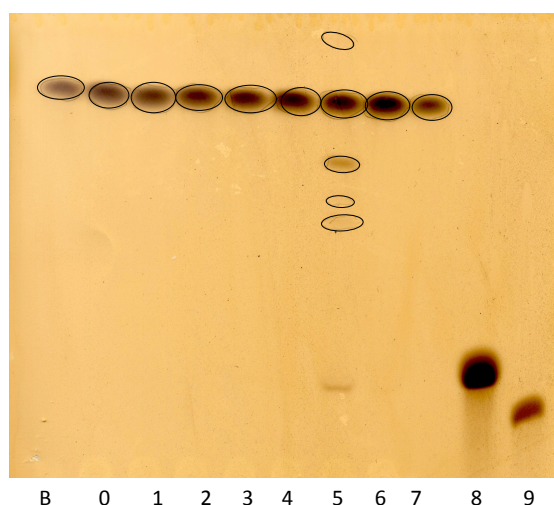


Figure 2.4.9: TLC analysis of enzymatic assays on 4Np-Glc (5 mM). The circle indicated the UV-visible products.

We analyzed the mode of action of GH10-XA by performing a time course incubation on 2Np-Cel. Reaction mixtures were analyzed by HPAEC-PAD revealing that the hydrolysis occurred between 2-nitrophenol and cellobiose with the accumulation of the latter after 16 h (Figure 2.4.10).

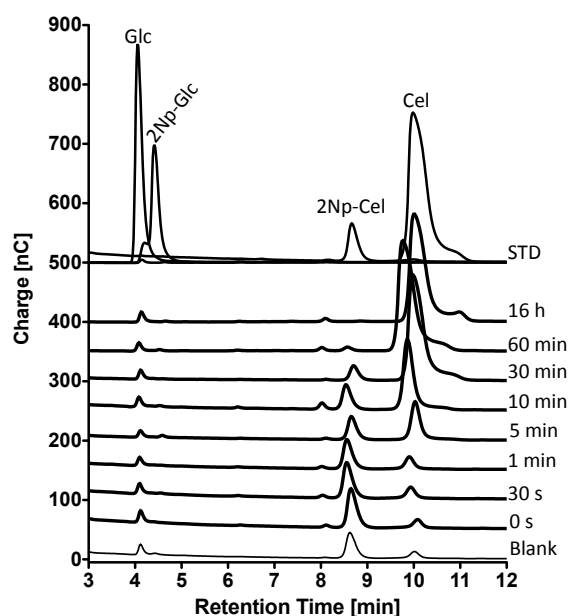


Figure 2.4.10: HPAEC-PAD analysis of the time course on 8 mM 2Np-Cel. Glc: Glucose; 2Np-Glc: 2Np-β-Glucoside; 2Np-Cel: 2Np-β-cellobioside; Cel: cellobiose.

In addition, GH10-XA hydrolyzed xylooligosaccharides (XOs) primarily to xylobiose and, less efficiently, to xylose after 16 h of incubation (Figure 2.4.11): These experiments indicated that GH10-XA removed disaccharidic units from the *non-reducing* end of the substrate as shown for other GH10 enzymes (Fontes CM *et al.*, 2000; Usui K *et al.*, 2003)

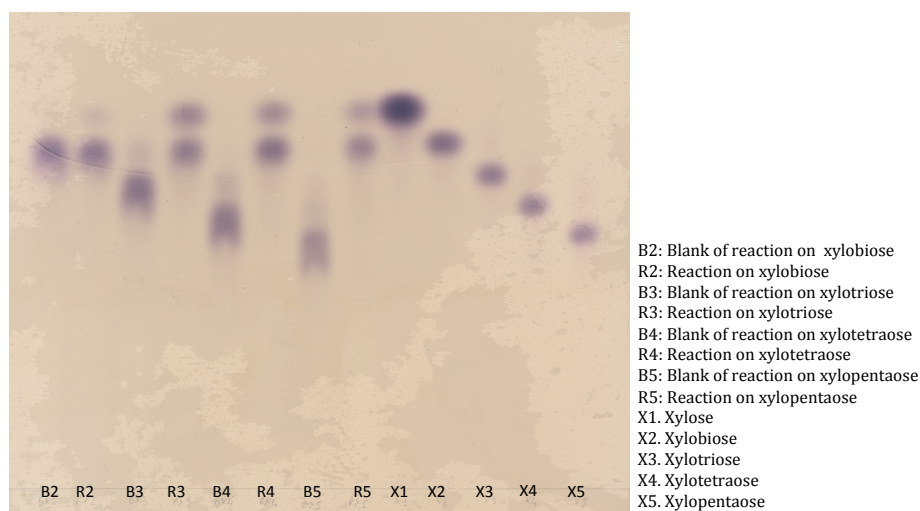


Figure 2.4.11: TLC analysis of the enzymatic assays on XOs (1 mg mL^{-1}). The reactions were carried out for 16 h at 65°C .

GH67-GA and GH67-GC

The hydrolytic activities of GH67-GA and GH67-GC were evaluated on MGX, beechwood xylan, aldouronic acids and 4Np- α -glucuronide (4Np-GlcUA) (Table 2.4.10). Both α -glucuronidases were not active on MGX and beechwood xylan but hydrolyzed aldouronic acids.

Substrate	U mg ⁻¹	
	GH67-GA	GH67-GC
Aldouronic acids	4.6 ± 0.3	2.6 ± 0.2
4Np-GlcUA	ND	0.09 ± 0.02
MGX	ND	ND
Beechwood xylan	ND	ND

Table 2.4.10: Assays of GH67-GA and GH67-GC on different substrates. ND: not detectable.

We analyzed by HPAEC-PAD the reaction products on aldouronic acids of GH67-GA and GH67-GC; the α -glucuronidase from *Geobacillus stearothermophilus* (AGUSB) was used as positive control. This analysis allowed the unequivocal identification of xylose (Xyl1), xylobiose (Xyl2), xylotriose (Xyl3), and xylotetraose (Xyl4) oligosaccharides, and of an additional peak eluting at about 15 min of the sodium acetate gradient (Figure 2.4.12). This peak can be ascribed to 4-*O*-methyl-glucuronic acids by comparing the positive control and others glucuronidases reported in literature (Tenkanen M *et al.*, 2000).

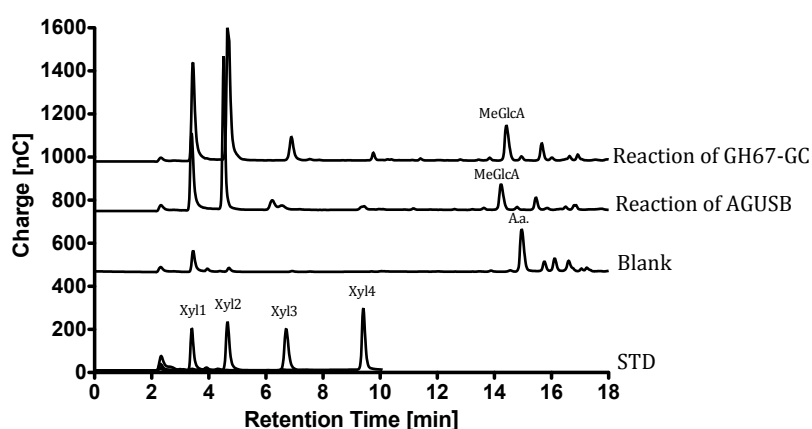


Figure 2.4.12: HPAEC-PAD analysis of GH67-GC reaction on aldouronic acids. The reaction of α -glucuronidase AGUSB was used as positive control. MeGlcA: 4-*O*-methylglucuronic acid; A.a.: aldouronic acids; xyl1: xylose; xyl2: xylobiose; xyl3: xylotriose; xyl4: xylotetraose.

The steady state kinetic constant of GH67-GA and GH67-GC on aldouronic acids and 4Np-GlcUA were measured in standard conditions and listed in Table 2.4.11.

Substrate	k_{cat} (s ⁻¹)		K_M (mg mL ⁻¹)		k_{cat}/K_M (s ⁻¹ mg ⁻¹ mL)	
	GH67-GA	GH67-GC	GH67-GA	GH67-GC	GH67-GA	GH67-GC
Aldouronic acids	11.7 ± 0.7	4.8 ± 0.3	0.23 ± 0.05	0.16 ± 0.03	51	30
4Np-GlcUA	ND	0.16 ± 0.03	ND	2.93 ± 0.86	ND	0.05

Table 2.4.11: Kinetic constants of GH67-GA and –GC on aldouronic acids and 4Np-GlcUA. The assays were performed in standard conditions.

The two enzymes have similar k_{cat}/K_M values for aldouronic acids while GH67-GC was also able to hydrolyze 4Np-GlcUA, albeit less efficiently. This is the only α -glucuronidase studied so far that is able to recognize as substrate both glucuronoxylan oligomers and 4Np-GlcUA. Indeed, most of the GH67 α -glucuronidases are only active on 4-*O*-MeGlcA linked to the *non-reducing* end of xyloligomers (Nurizzo D *et al.*, 2002; Golan G *et al.*, 2004; Jia XJ *et al.*, 2014) and only few studies demonstrated their high activity against longer polymeric substrates (Tenkanen M *et al.*, 2000; Ryabova O *et al.*, 2009). By contrast, the Agu4A α -glucuronidase from *Thermotoga maritima* from family GH4, which efficiently hydrolysed 4Np-GlcUA, was unable to recognize glucuronoxylans as substrates (Suresh C *et al.*, 2003; Suresh C *et al.*, 2002).

Synergic action of GH10 and GH67 on glucuronoxylans

Both 4-*O*-methyl-glucuronoxylan and beechwood xylan were branched by 4-*O*-methyl- α -glucuronide residues bound to the xylan backbone. The substrate specificity of the GH10 and GH67 enzymes prompted us to analyse the combined effect of the xylanase and of the α -glucuronidases on these substrates. Preliminary analyses were performed in 100 mM sodium phosphate buffer, pH 6.0 at 50 ° C in the presence of MGX (5 mg mL⁻¹) with 0.26 μ M of GH10-XA and

0.11 μM of GH67-GA for 16 h. The hydrolysis products were measured by Somogyi-Nelson assay. These tests showed that the presence of GH67-GA slightly increased the activity of the xylanase, leading to an increment of the reducing ends of 2 mM after 16 h (Table 2.4.12); however, GH67-GA showed a low stability at these operational conditions, therefore we did not tested it any further.

Enzymes	Products (mM)
GH10-XA	4.0 ± 0.4
GH67-GA	ND
GH10-XA + GH67-GA	6.0 ± 0.6

Table 2.4.12: Measures of reducing ends produced by the combined action of GH10-XA and GH67-GA on 5 mg mL⁻¹ MGX.

Instead, the combined action of GH10-XA (0.26 μM) and of GH67-GC (0.09 μM) was tested on both MGX (5 mg mL⁻¹) and beechwood xylan (15 mg mL⁻¹) in buffer sodium phosphate 100 mM pH 6.5 at 65 °C. After only 2 h of incubation we measured an increment of the reducing ends of 1 mM on MGX and 2 mM on beechwood xylan (Table 2.4.13).

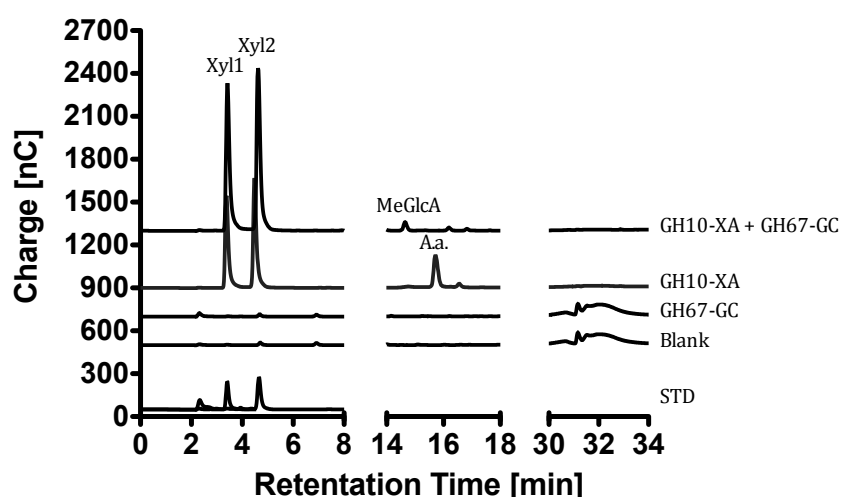
MGX	
Enzymes	Products (mM)
GH10-XA	8.6 ± 0.4
GH67-GC	ND
GH10-XA + GH67-GC	9.4 ± 0.1
Beechwood xylan	
GH10-XA	24.8 ± 0.6
GH67-GC	ND
GH10-XA + GH67-GC	26.7 ± 0.3

Table 2.4.13: Measures of reducing ends produced by the combined action of GH10-XA and GH67-GC on 5 mg mL⁻¹ MGX and 15 mg mL⁻¹ beechwood xylan.

We compared the combined action of GH10-XA and GH67-GC to their GH10 and GH67 homologs from *C. lactoaceticus* (Jia XJ *et al.*, 2014). Both reactions produced the same amounts of reducing ends from beechwood xylan but GH10-XA and GH67-GC obtained this results much quicker (after 2 h vs 12 h). In addition, to achieve these amounts of products we used much less enzyme concentration (7- and 9-fold for GH10 and GH67 enzymes, respectively) than the one used for *C. lactoaceticus* enzymes.

We analyzed these reaction products by HPAEC-PAD confirming that the enzymes clearly released xylose, xylobiose, and 4-*O*-methyl-glucuronic acid: the direct identification, for the first time, of the latter compound confirmed that GH10-XA and GH67-GC cooperate in the hydrolysis of beechwood xylan (Figure 2.4.13). The hydrolysis catalyzed by GH10-XA alone led to accumulation of xylobiose, and a third peak, retained more time in the column than 4-*O*-methyl-glucuronic acids and presumably corresponding to aldouronic acids (Figure 2.4.13).

A



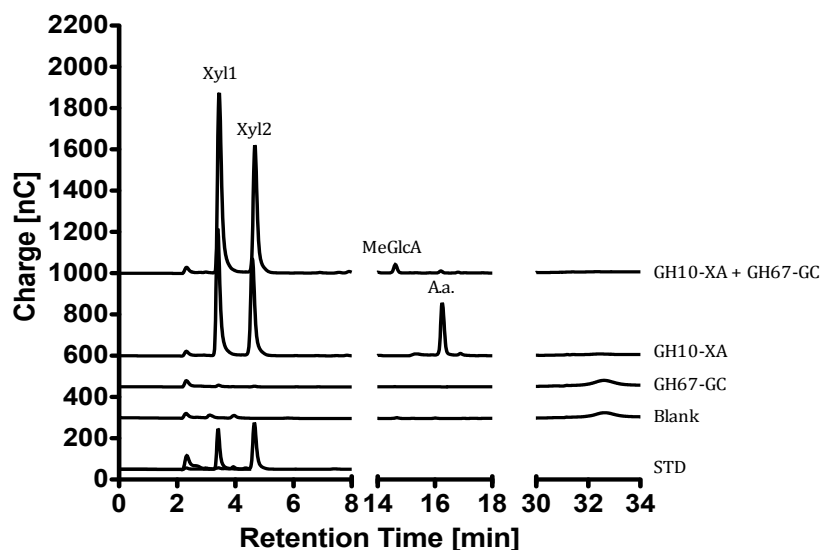
B

Figure 2.4.13: HPAEC-PAD analysis of the combined effects of GH10-XA and GH67-GC on MGX (A) and beechwood xylan (B). Xyl1: xylose; xyl2: xylobiose; MeGlcA: 4-O-methylglucuronic acid; Aa.: aldouronic acids.

Combined action of GH10-XA, GH67-GC and GH3-XT on glucuronoxylans

On the basis of these encouraging results, and to further improve the biotransformation reactions catalyzed by GH10-XA and GH67-GC, we selected in CAZy database the β -xylosidase (GH3-XT) from *Thermotoga thermarum*, classified in family GH3 (Shi H *et al.*, 2013) as a third catalyst to include in the enzymatic cocktail. This enzyme displayed extraordinary and desirable properties: high expression levels, high xylose tolerance, thermal stability (it maintained 90% activity after 120 min at 75°C), and excellent activity on xylooligosaccharides (Shi H *et al.*, 2013). We added GH3-XT to the enzymatic mixture and we carried out the reaction on MGX and beechwood xylan. Interestingly, the reducing ends on both substrates increased substantially in presence of the three enzymes after 2 h of incubation (Table 2.4.14).

MGX	
Enzymes	Products (mM)
GH10-XA	8.6 ± 0.4
GH67-GC	ND
GH3-XT	2.1 ± 0.1
GH10-XA + GH67-GC	9.4 ± 0.1
GH10-XA + GH67-GC + GH3-XT	14.3 ± 0.9
Beechwood xylan	
GH10-XA	24.8 ± 0.6
GH67-GC	ND
GH3-XT	1.1 ± 0.2
GH10-XA + GH67-GC	26.7 ± 0.3
GH10-XA + GH67-GC + GH3-XT	33.9 ± 0.2

Table 2.4.14: Reducing ends amounts produced by the combined action of GH10-XA, GH67-GC, and GH3-XT on 5 mg mL⁻¹ MGX and 15 mg mL⁻¹ beechwood xylan.

To assess the potential use of the three enzymes in saccharification of plant biomasses we analysed the activities of GH10-XA, GH67-GC and GH3-XT on the liquid fraction of *A. donax* pretreated by steam-explosion. The analysis of the reaction mixtures by HPAEC-PAD revealed the production of xylose as the final product demonstrating that the three enzymes were active on this biomass (Fig. 2.4.14). The absence of 4-O-methyl-glucuronic acid, the other expected final product (Fig. 2.4.14) could be due to the low amounts of this substituent in the pretreated biomass.

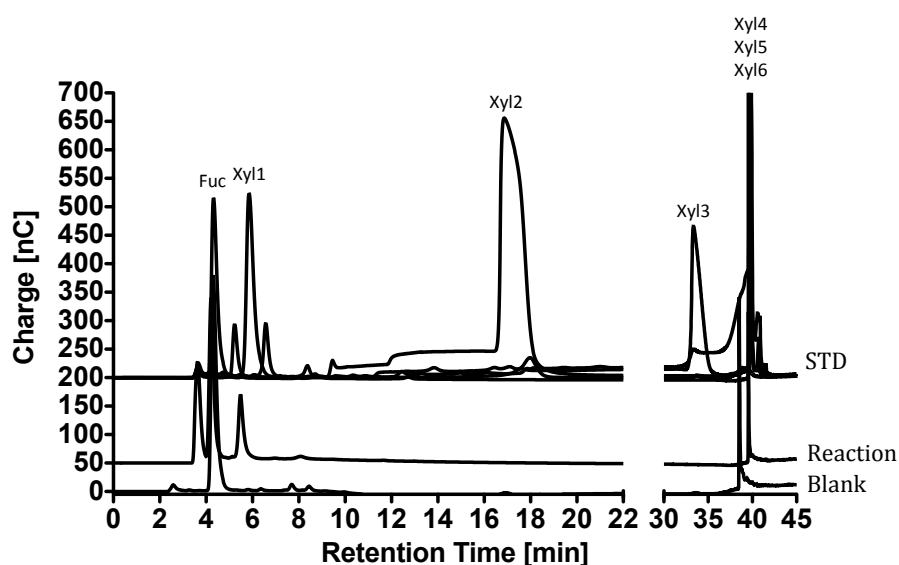


Figure 2.4.14: HPAEC-PAD analysis of enzymatic saccharification performed by GH10-XA, GH67-GC, and GH67-XT on *Arundo donax* pretreated. Fuc: internal standard fucose; xyl1: xylose; xyl2: xylobiose; xyl3: xylotriose; xyl4: xylotetraose; xyl5: xylopentaose.

The activity of this enzymatic cocktail on a pretreated biomass coming from a biorefinery production plant is of particular importance as it demonstrates that the use of enzymes from thermophiles is very close to practical application and makes them promising candidates for the treatment of plant biomasses.

2.5- References

- Aditiya HB, Chong WT, Mahlia TM, Sebayang AH, Berawi MA, Nur H. Second generation bioethanol potential from selected Malaysia's biodiversity biomasses: A review. *Waste Manag.* 2016 Jan;47(Pt A):46-61.
- Bai Y, Wang J, Zhang Z, Yang P, Shi P, Luo H, Meng K, Huang H, Yao B. A new xylanase from thermoacidophilic *Alicyclobacillus* sp. A4 with broad-range pH activity and pH stability, *J. Indus. Microbiol. Biotechnol.* 37 (2010) 187–194.
- Bhardwaj A, Mahanta P, Ramakumar S, Ghosh A, Leelavathi S, Reddy VS. Emerging role of N- and C-terminal interactions in stabilizing (β/α)₈ fold with special emphasis on Family 10 xylanases *Comput Struct Biotechnol J.* 2012 Nov 1;2:e201209014.
- Biely P, Vršanská M, Tenkanen M, Kluepfel D (1997) Endo- β -1,4-xylanases families: differences in catalytic properties. *J Biotechnol.* 57: 151–166
- Blanco A, Diaz P, Martinez J, Lopez O, Soler C, Pastor FI. (1996) Cloning of a *Bacillus* sp. BP-23 gene encoding a xylanase with high activity against aryl xylosides, *FEMS Microbiol. Lett.* 137 1996 285–290.
- Cheng, J.J., Timilsina, G.R., 2011. Status and barriers of advanced biofuel technologies: a review. *Renew. Energy* 36 (12), 3541–3549.
- Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M, Identification of an archaeal alpha-l-fucosidase encoded by an interrupted gene. Production of a functional enzyme by mutations mimicking programmed -1 frameshifting (vol 278, pg 14,622, 2003), *J. Biol. Chem* 278 (2003) 47350.
- Collins T, Gerday C, Feller G, Xylanases, xylanase families and extremophilic xylanases, *FEMS Microbiol. Rev.* 29 (2005) 3–23.
- Collins T, Meuwist M-A, Stals I, Claeysens M, Fellert G, Gerday C (2002) A novel family 8 xylanase, functional and physicochemical characterization. *J Biol Chem* 277: 35133–35139
- Correia MA, Mazumder K, Brás JL, Firbank SJ, Zhu Y, Lewis RJ, York WS, Fontes CM, Gilbert HJ. (2011) Structure and function of an arabinoxylan-specific xylanase. *J Biol Chem* 286: 22510–22520
- Di Nasso NNO, Roncucci N, Bonari E (2013) Seasonal dynamics of aboveground and belowground biomass and nutrient accumulation and

remobilization in giant reed (*Arundo donax* L.): a three-year study on marginal land. *Bioenergy Res* 6(2):725–736.

- Diodato N, Fagnano M, Alberico I (2009) CliFEM – climate forcing and erosion response modelling at long-term sele river research basin (Southern Italy). *Nat Hazards Earth Syst Sci* 9:1693–1702.
- Du XL, Jiang Z, Su DS, Wang JQ. Research Progress on the Indirect Hydrogenation of Carbon Dioxide to Methanol. *ChemSusChem*. 2016 Feb;9(4):322-32.
- European Commission, 2012. Proposal for a Directive of the European Parliament and of the Council amending Directive 98/70/EC relating to the quality of petrol and diesel fuels and amending Directive 2009/28/EC on the promotion of the use of energy from renewable sources, Brussels, 17.10.2012 SWD(2012) 343 final. Available at: <http://ec.europa.eu/>, Accessed 18 January 2015).
- European Commission, 2015. Directive (EU)2015/1513 of the European Parliament and of the Council of 9 September 2015 amending Directive 98/70/EC relating to the quality of petrol and diesel fuels and amending Directive 2009/28/EC on the promotion of the use of energy from renewable sources. Official Journal of the European Union, L 239/1 of 15.09.2015 Available at: <http://eur-lex.europa.eu> (Accessed 8 November 2015).
- Fiorentino N, Impagliazzo A, Ventorino V, Pepe O, Piccolo A, Fagnano M (2010) Biomass accumulation and heavy metal uptake of giant reed on polluted soil in southern Italy. *J Biotechnol* 150:261
- Fontes CM, Gilbert HJ, Hazlewood GP, Clarke JH, Prates JA, McKie VA, Nagy T, Fernandes TH, Ferreira LM. A novel *Cellvibrio mixtus* family 10 xylanase that is both intracellular and expressed under non-inducing conditions, *Microbiology* 146 (Pt (8)) (2000) 1959–1967.
- Gilkes NR, Claeysens M, Aebersold R, Henrissat B, Meinke A, Morrisson HD Kilburn DG, Warren RA, Miller RC Jr. Structural and functional relationships in two families of beta-1,4-glycanases (1991). *Eur J Biochem* 202: 367–377
- Golan G, Shallom D, Teplitsky A, Zaide G, Shulami S, Baasov T, Stojanoff V, Thompson A, Shoham Y, Shoham G. Crystal structures of *Geobacillus*

- stearothermophilus alpha-glucuronidase complexed with its substrate and products - Mechanistic implications, *J. Biol. Chem.* 279 (2004) 3014–3024.
- Jia XJ, Mi SF, Wang JZ, Qiao WB, Peng XW, Han XJ, Insight into glycoside hydrolases for debranched xylan degradation from extremely thermophilic bacterium *Caldicellulosiruptor lactoaceticus*, *PLoS One*. 2014 Sep 3;9(9):e106482
 - Kim DY, Shin DH, Jung S, Kim H, Lee JS, Cho HY, Bae KS, Sung CK, Rhee YH, Son KH, Park HY. Novel Alkali-Tolerant GH10 Endo-beta-1,4-Xylanase with broad substrate specificity from microbacterium *trichothecenolyticum* HY-17, a gut bacterium of the mole cricket *Gryllotalpa orientalis*, *J. Microbiol. Biotechnol.* 24 (2014) 943–953.
 - Kim S, Dale BE Global potential bioethanol production from wasted crops and crop residues. *Biomass and Bioenergy* 26 (2004) 361–375.
 - Kolenova K, Vrsanska M, Biely P, Mode of action of endo-beta-1,4-xylanases of families 10 and 11 on acidic xylooligosaccharides, *J. Biotechnol.* 121 (2006) 338–345.
 - Krengel U, Dijkstra BW (1996) Three-dimensional structure of Endo-1,4-beta-xylanase I from *Aspergillus niger*, molecular basis for its low pH optimum. *J. Mol. Biol.* 263: 70–78
 - Lagaert S, Pollet A, Courtin CM, Volckaert G. β -Xylosidases and α -L-arabinofuranosidases: Accessory enzymes for arabinoxylan degradation. *Biotechnology Advances* 32 (2014) 316–332
 - Lee YE, Zeikus JG. Genetic organization, sequence and biochemical characterization of recombinant β -xylosidase from *Thermoanaerobacterium saccharolyticum* strain B6A-RI. *J Gen Microbiol* 1993;139:1235–43.
 - Lennartsson PR , Erlandsson P, Taherzadeh MJ Integration of the first and second generation bioethanol processes and the importance of by-products. *Bioresource Technology* 165 (2014) 3–8
 - Lin, Y., Tanaka, S., 2006. Ethanol fermentation from biomass resources: current state and prospects. *Appl. Microbiol. Biotechnol.* 69 (6), 627–642.
 - Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B, The carbohydrate-active enzymes database (CAZy) in 2013, *Nucleic Acids Res.* 42 (2014) D490–5.

- Lombardi, M., Rana, R., Tricase, C., Ingrao, C., 2014. Sustainability criteria and certification schemes of biofuels in the European Union. In: Adamczyk, W. (Ed.), *Commodity Science in Research and Practice - Towards sustainable development*. Foundation of the Cracow University of Economics, Cracow, pp. 127–137.
- Madhavan, A., Srivastava, A., Kondo, A., Bisaria, V.S., 2012. Bioconversion of lignocellulose-derived sugars to ethanol by engineered *Saccharomyces cerevisiae*. *Crit. Rev. Biotechnol.* 32 (1), 22–48.
- Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y.Y., Holtzapple, M., Ladisch, M., 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour. Technol.* 96 (6), 673–686.
- Nurizzo D, Nagy T, Gilbert HJ, Davies GJ, The structural basis for catalysis and specificity of the *Pseudomonas cellulosa* alpha-glucuronidase, GlcA67A, *Structure* 10 (2002) 547–556.
- Østergård H, Hauggaard-Nielsen H, and Pilegaard K, Bioenergy efficiency improvements. DTU International Energy Report 2012
- Peng X, Qiao W, Mi S, Jia X, Su H, Han Y. Characterization of hemicellulase and cellulase from the extremely thermophilic bacterium *Caldicellulosiruptor owensensis* and their potential application for bioconversion of lignocellulosic biomass without pretreatment. *Biotechnol Biofuels*. 2015 Aug 28;8:131.
- Rana R, Ingrao C, Lombardi M, Tricase C. Greenhouse gas emissions of an agro-biogas energy system: Estimation under the Renewable Energy Directive *Science of the Total Environment* 550 (2016) 1182–1195
- REN21, 2012. Renewables 2012 Global Status Report. REN21 Secretariat, Paris.
- Ruile P, Winterhalter C, and Liebl W. Isolation and analysis of a gene encoding alpha-glucuronidase, an enzyme with a novel primary structure involved in the breakdown of xylan. *Mol Microbiol.* 1997 Jan;23(2):267-79.
- Ryabova O, Vrsanska M, Kaneko S, van Zyl VH, Biely P, A novel family of hemicellulolytic alpha-glucuronidase, *FEBS Lett.* 583 (2009) 1457–1462.

- Shi H, Li X, Gu HX, Zhang Y, Huang YJ, Wang L, Wang F., Biochemical properties of a novel thermostable and highly xylose-tolerant beta-xylosidase/alpha-arabinosidase from *Thermotoga thermarum*, *Biotechnol Biofuels*. 2013 Feb 20;6(1):27.
- Somogyi M (1952) Estimation of sugars by colorimetric method. *J Biol Chem* 200: 245.
- St. John FJ, Rice JD, Preston JF (2006) Characterization of XynC from *Bacillus subtilis* subsp. *subtilis* strain 168 and analysis of its role in depolymerization of glucuronoxylan. *J Bacteriol* 188: 8617–8626
- Suresh C, Rus'd AA, Kitaoka M, Hayashi K, Evidence that the putative alpha-glucosidase of *Thermotoga maritima* MSB8 is a pNP alpha-d-glucuronopyranoside hydrolyzing alpha-glucuronidase, *FEBS Lett.* 517 (2002) 159–162.
- Suresh C, Kitaoka M, Hayashi K, A thermostable non-xylanolytic alpha-glucuronidase of *Thermotoga maritima* MSB8, *Biosci. Biotechnol. Biochem.* 67 (2003) 2359–2364.
- Tenkanen M, Siika-aho M, An alpha-glucuronidase of *Schizophyllum commune* acting on polymeric xylan, *J. Biotechnol.* 78 (2000) 149–161.
- Thompson J, Pikis A, Ruvinov SB, Henrissat B, Yamamoto H, and Sekiguchi J. The gene *glvA* of *Bacillus subtilis* 168 encodes a metal-requiring, NAD(H)-dependent 6-phospho-alpha-glucosidase. Assignment to family 4 of the glycosylhydrolase superfamily. *J Biol Chem.* 1998
- Uday US, Choudhury P, Bandyopadhyay TK, Bhunia B. Classification, mode of action and production strategy of xylanase and its application for biofuel production from water hyacinth. *Int J Biol Macromol.* 2016 Jan;82:1041-54.
- Usui K, Suzuki T, Akisaka T, Kawai K, A cytoplasmic xylanase (XynX) of *Aeromonas caviae* ME-1 is released from the cytoplasm to the periplasm by osmotic downshock, *J. Biosci. Bioeng.* 95 (2003) 488–495.
- Van Petegem F, Collins T, Meuwist M-A, Gerday C, Feller G, Van Beeumen J (2003) The structure of a cold-adapted family 8 xylanase at 1.3 Å resolution. *J Biol Chem* 278: 7531–7539

- Vimala Rodhe A, Sateesh L, Sridevi J, Venkateswarlu B, Venkateswar Rao L. Enzymatic hydrolysis of sorghum straw using native cellulase produced by *T. reesei* NCIM 992 under solid state fermentation using rice straw. 3 Biotech. 2011 Dec;1(4):207-215.
- Vršanská M, Kolenová K, Puchart V, Biely P (2007) Mode of action of glycoside hydrolase family 5 glucuronoxylan xylanohydrolase from *Erwinia chrysanthemi*. FEBS J 274: 1666–1677.

3 ~ Temporal changes of microbial community in Pisciarelli Solfatara

Chapter 3 describes the second part of my PhD project consisting in the weekly monitoring of the microbial community and of the geochemical phenomena occurred in the hot spring Pisciarelli Solfatara (Naples, Italy). This section begins with a brief introduction on the populations inhabiting hot springs in the world (paragraph 3.1.1) and the metagenomic study conducted by our lab in Pisciarelli (paragraph 3.1.2). The paragraph 3.2 describes the purpose of this work, while the experimental part is shown in paragraph 3.3. Results and discussion are described in paragraph 3.4. At last, the paragraph 3.5 reports the literature references included in this chapter.

3.1- Introduction

3.1.1- Metagenomic of hyperthermophilic environments

Metagenome sequencing, coupled with transcriptomics, proteomics and metabolomics, has led to the development of sophisticated systems biology approaches (Zhang W *et al.*, 2010), which facilitate the combined study of the functional relationships within the microbial community and their interaction with the environment (Cowan DA *et al.*, 2015). The development of SSU rRNA phylogenetics revealed the high complexity and diversity of the prokaryotic phylotypes across the thermophilic biotopes on Earth (spanning a pH gradient from 0 to >10, and a temperature gradients from 60 °C to >120 °C). Sites of volcanic activity all over the land and under the sea provide a variety of different environments for extremophilic Archaea and Bacteria. Hot springs populated by hyperthermophiles are very diverse and some of them show combinations of extreme conditions, such as, acidic, alkaline, high salinity or pressure, high concentrations of heavy metals (Cowan DA *et al.*, 2015). A large number of hot springs is very close to each others, for instance Yellowstone National Park includes more than 10.000 thermal sites such as geysers, mud pools, hot springs and vents. Well-known examples of such extreme environments are terrestrial surface hot springs. This diversity of habitats provides a vast number of sites to sample, all with potential interest for metagenomic analysis. The high temperatures in hot springs, which is usually over the limit of eukaryotic life (near to 60 °C), allow growths of only Bacteria and Archaea (and their viruses). In addition, high temperature is one of the main factors that shapes microbial communities in hot springs during environmental or geochemical changes occurring in time (Miller SR *et al.*, 2009). A particular example is a unique coastal hot springs in Reykjanes peninsula (Iceland), where the microbial populations are exposed to fluctuations of temperature (45 to 95°C) as a consequence of periodic high tides. The areas showing the longest (6 h) hot-temperature periods showed a majority of terrestrial thermophilic Bacteria,

whereas areas with the shortest (2 h) hot-temperature periods showed predominance of moderately thermophilic microorganisms and the presence of mesophilic marine microorganisms (Hobel, CFV *et al.*, 2005). In addition to temperature as main determinant, the structure and the diversity of the populations living in hot springs may be subjected also to geochemical features (Meyer-Dombard DR *et al.*, 2005).

A wider metagenomic survey of the microbial species in the Yellowstone National Park (YNP) was addressed by Inskeep and co-workers in 2010 (Inskeep WP *et al.* 2010) aiming to identify the predominant microbial populations living in five geochemically different high-temperature environments. Inskeep and collaborators observed that high-temperature springs with acidic pH were dominated by unknown Archaea, which are distantly related to organisms whose genomes have been sequenced corresponding, respectively, to the Order Sulfolobales (in Crater Hills, 75 °C, pH 2.5 and Norris Geyser Basin, 65 °C, pH 3.0) and Thermoproteales (in Joseph's Coat Hot Spring, 80 °C, pH 6.1). Conversely, the two microbial communities inhabiting Calcite (75 °C) and Mammoth Hot Springs (71 °C) with pH above 6.0 were dominated by bacterium *Hydrogenobaculum* belonging to the Order Aquificales (Inskeep WP *et al.*, 2010).

More recently, Menzel and coworkers conducted a metagenomic study on eight globally distributed terrestrial hot springs from Eryuan (China), Grensdalur and Krìsuvík (Iceland), Pisciarelli and Pozzuoli (Italy), Uzon Caldera (Russia), and the YNP (USA) with a temperature range between 61 °C and 92 °C and pH between 1.8 and 7.0. Archaeal species that are found in all eight samples are *Thermoplasma volcanium*, *Ferroplasma acidarmanus*, *Sulfolobus tokodaii*, *Sulfolobus acidocaldarius*, *Sulfolobus solfataricus*, *Sulfolobus islandicus*, *Metallosphaera yellowstonensis*, *Metallosphaera sedula*, and *Acidianus hospitalis*. While the first two belong to Thermoplasmatales, the rest belongs to the family

Sulfolobaceae. No single bacterial species common for all eight samples was found, however, the bacterial species found in six of the eight samples (except Pozzuoli and Krìsuvík) are *Ammonifex degensii*, *Salinibacter ruber*, *Sulfurihydrogenibium azorense*, *Hydrogenobaculum* sp. Y04AAS1, *Hydrogenobacter thermophilus* and *Aquifex aeolicus*. Except for the first two, all species belong to the order Aquificales (Menzel P *et al.*, 2015).

In 2014, Wang and colleagues reported the correlation between microbial diversity and geochemistry by considering the seasonal changes (Wang S *et al.*, 2014). The study was conducted in several hot springs (temperature range 55-95 °C; pH range 1.9-9.2) in Tengchong (China), a subtropical area located on the northeastern edge of Tibet-Yunnan with heavy temporal monsoon rainfall (Briggs BR *et al.*, 2013). Wang and co-workers compared the samples collected in the rainy season (June and August) with samples collected in the dry season (January). They found that the seasonal effects on the microbial diversity are more pronounced in sediments relative to water samples. In acidic springs the water communities between January and June were highly similar to each other and both were predominated by crenarchaeon *Sulfolobus*. Instead, in August, the bacterium *Hydrogenobacter* became the most abundant taxon followed by *Sulfolobus*. In the sediments of two acidic springs from January to August the dominant archaeal taxon *Sulfolobus* was replaced by bacteria *Desulfococcus* and archaeon *Ignisphaera*. In the neutral-alkaline spring, the water community remained constant in both dry and rainy seasons, except for the Shuirebaozha spring where, among the most abundant bacterial taxa, *Fervidobacterium* changed in *Hydrogenobacter* from June to August. The two alkaline springs, characterized by fast-flowing and high discharges, harbored *Hydrogenobacter* in January and June but the community structures showed a notable change in August where *Persephonella* and the candidate phylum OP1 became the dominant members (Wang S *et al.*, 2014).

Wang and co-workers suggested that pH was a primary factor influencing the microbial community shifts, followed by temperature and dissolved organic carbon (DOC), but the exact reasons of the seasonal changes in microbial community structure are still unknown.

3.1.2 Metagenomic analysis of Pisciarelli Solfatara

Metagenomics of extreme environments play a key role in the discovery of new enzymes with unique features crucial for industrial applications, such as the enzymatic hydrolysis of (hemi)cellulose for the efficient production of second generation bioethanol (Chapters 2 of this thesis). The screening for novel biocatalysts from metagenomic datasets of extreme environments represents a valuable alternative to time consuming classical screenings of isolated cell cultures and complex protein engineering procedures for the optimization of available enzymes from mesophiles. In the framework of these considerations, in our laboratory we embarked in the study, through a metagenomic approach, of the communities of hyperthermophiles populating the volcanic site Pisciarelli Solfatara to understand their evolution and adaptation to extreme conditions and to obtain a rich source of biocatalysts.

Solfatara volcano is located in the central part of the repeatedly collapsed Phlegraean Fields caldera (diameter 12–15 km), and is one of the youngest volcanoes formed within this active volcanic field (Rosi M and Sbrana S, 1987; Orsi G *et al.* 1996; Isaia R *et al.* 2009). (Figure 3.1.1).

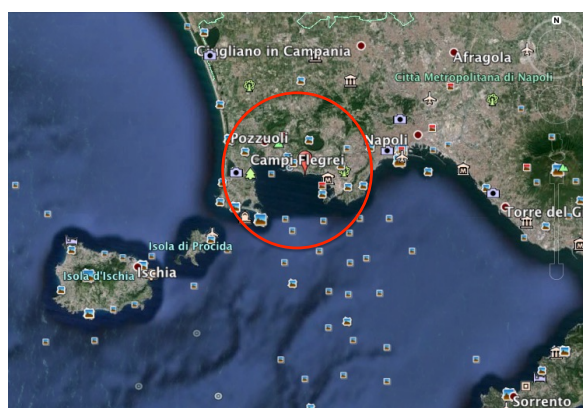


Figure 3.1.1: Phlegraean Fields caldera

Pisciarelli Solfatara is only about 800 m² in size and has two separated mud pools, Pool1 and Pool2 (Figure 3.1.2). Although the Pools are very close, they

show different physical characteristic: indeed, the temperatures measured in Pool1 and Pool2 were 86 °C and 92 °C, while the pHs were 5.5 and 1.5 respectively.

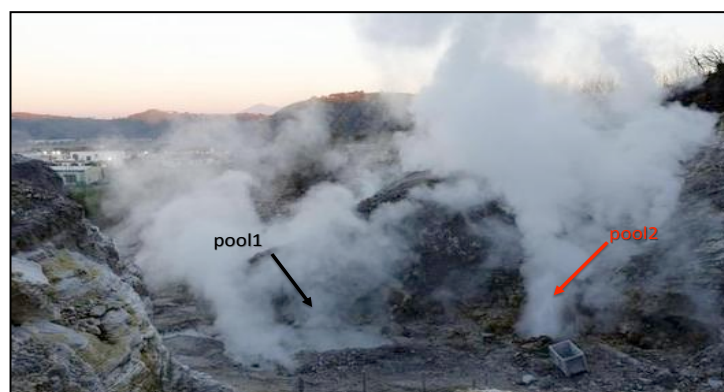


Figure 3.1.2: Pisciarelli hot springs.

In 2013, in our laboratory the metagenomic analyses on Pisciarelli were begun: the extracted metagenomic DNA samples from Pool1 and Pool2, respectively, was sequenced by Illumina PE technology, in collaboration with Dr Yizhuang Zhou and Dr Jin Xu (BGI, Hong Kong), obtaining 33,240,694 and 26,507,584 reads for Pool1 and Pool2, respectively. The searches in database revealed that for Pool1 and Pool2, respectively, 32 % and 62 % of the reads did not match with any sequences annotated in database, suggesting that these sequences derived from unknown species. All remaining reads (68% in Pool1 and 38% in Pool2) found matching with archaeal sequences, indicating that both pools are colonized mainly by Archaea and that the presence of hyperthermophilic bacteria was negligible ([Figure 3.1.3](#))

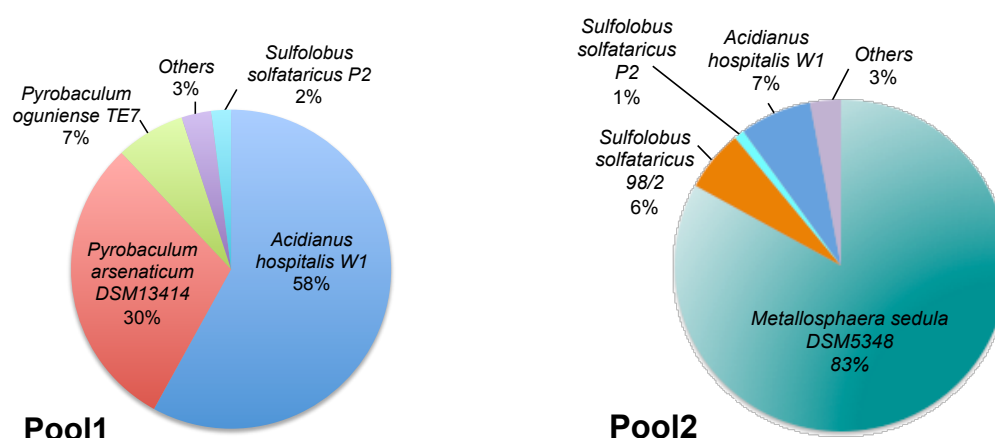


Figure 3.1.3: Charts of Pisciarelli species

Among the matching reads, Pool1 is dominated by the Genus *Acidianus*, followed by *Pyrobaculum* and with only traces of *Sulfolobus*, while Pool2 is colonized mainly by *Metallosphaera* following by *Acidianus hospitalis* and *Sulfolobus solfataricus*.

Successively, the reads were assembled in scaffolds which were analyzed by softwares Glimmer MG and Prodigal to predict the putative ORFs. These analyses revealed 14,934 and 17,652 ORFs in Pool1 and Pool2, respectively. They were analyzed by KEGG database, that allowed to identify a large number of ORFs involved in carbohydrate metabolism (not shown). Thus, the CAZy families present in Pool1 and Pool2 were identified in collaboration with Dr Bernard Henrissat (CNRS-Marseilles). Surprisingly, a large number of cazyme families involved in starch and hemicellulose hydrolysis were found in both pools (Figure 3.1.4), as well as enzymes involved in protein *N*-glycosylation.

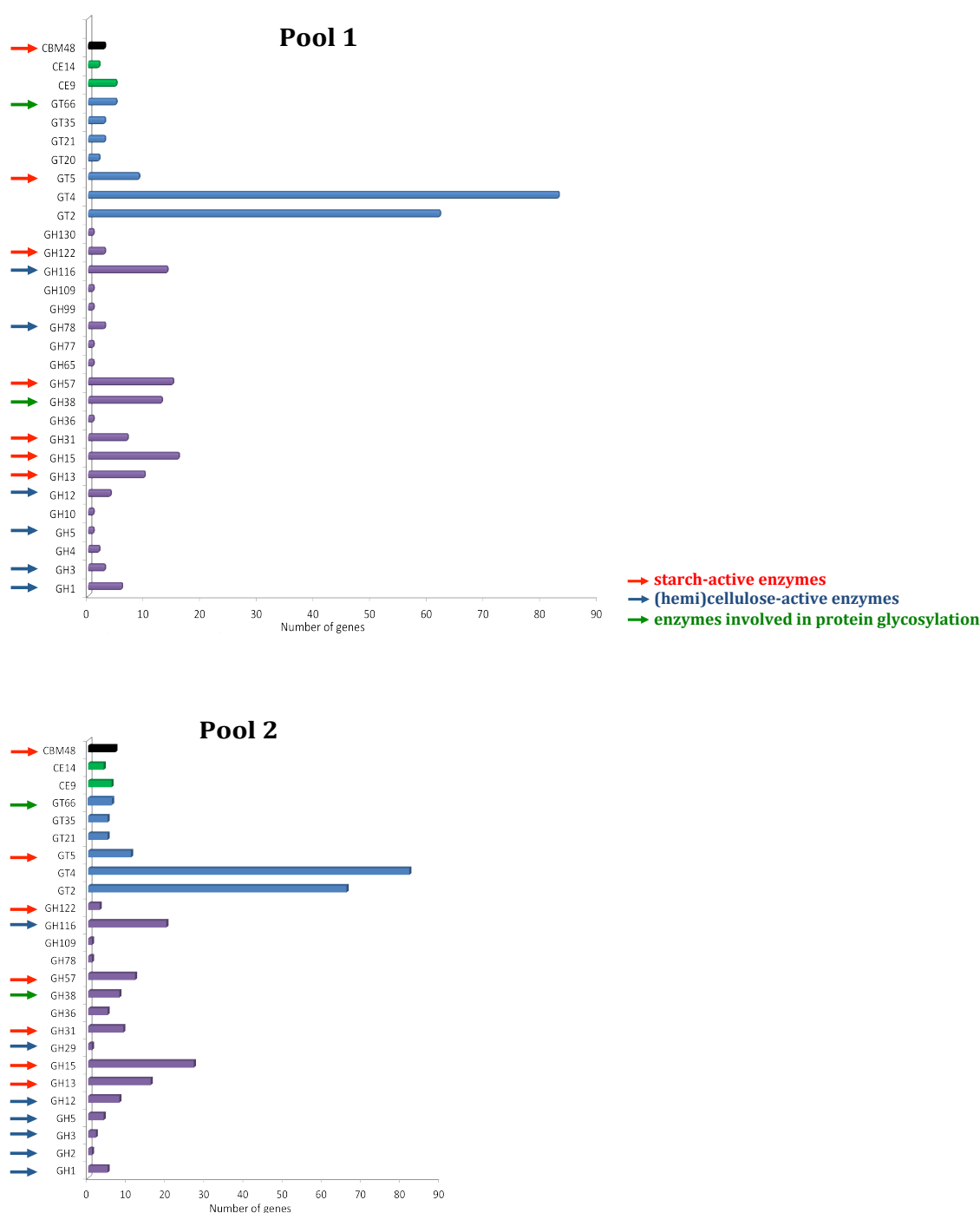


Figure 3.1.4: Bar Charts of identified ORF

These results suggested that the large amount of plants that surround the Solfatara could provide a carbon source on which the microorganisms grow. Indeed, rain and wind lead parts of these plants within the Pools, as shown in the [Figure 3.1.5](#), entering in close contact with the microorganisms.



Figure 3.1.5: Picture of the plants in Pool1

Thus, Pisciarelli Solfatara proved to be an excellent source of thermophilic CAZymes.

In last few years the geothermal activity of Pisciarelli has experienced a remarkable increase. Starting from 2003, this hot spring has been marked by a sequence of temperature peaks, by the opening of new vigorous vents and degassing pools, also accompanied by intense local seismic activity. (Fedele A *et al.*, 2014). These geothermal changes could be reflected in microbial diversity changes, affecting the number of species present in this environment. Therefore, microbial ecology analyses in Pisciarelli might emphasize which of the species present are more able to adapt to the rapid changes, as well as bring out new species and expand our knowledge on the metabolism of extremophiles. All this could increase the possibilities of identifying novel ORFs coding for interesting thermophilic CAZymes.

3.2 – Purpose of this study

The sources of new enzymes were technically limited to a minor fraction of total microbial diversity, the culturable microorganisms, which have been estimated as representing less than 1% of the real diversity in most environments (Amann RI *et al.* 1995). This drawback is even more prominent with extremophilic microorganisms. Thus, metagenomic studies of extreme environment could allow to identify novel interesting enzymes applicable in industrial processes, as the second generation bioethanol production. In addition, metagenomic studies could provide important informations on the microbiome inhabiting extreme environments. Studies on the microbial community structures have been performed in diverse hot springs or in different sites of the same hot spring (Inskeep WP *et al.*, 2010; Hou W *et al.*, 2013) but only a limited number of studies in the last decade have explored temporal changes in hot spring communities (Wang S *et al.*, 2014). In framework with these considerations, this part of my PhD project aims of perform a metagenomic study on the hot spring Pisciarelli (Naples) to determinate the changes of microbial community related to temporal and geochemical variations. The Pisciarelli microbiome was monitored for a year by weekly metagenomic DNA extractions. The DNA analysis of some samples has proven remarkable changes of the microbial community related with geochemical variations occurred in Pisciarelli in 2015.

3.3- Experimental Procedures

Sampling in Pisciarelli

Every week 2 L of water containing mud from Pool1 were withdrawn and the temperature was measured *in situ*. The mud pH was measured in lab at temperature of 40-45 °C. An aliquot of sample (about 100 mL) was filtered (*cut-off* 0.45 µm and then 0.22 µm) for water composition analyses.

Metagenomic DNA extraction

The water containing mud was centrifuged at 6000 x g for 30 min at 20 °C. The recovered sediment was conserved at -20 °C. The metagenomic DNA was extracted following the protocol reported in Zhou *et al.*, 1996 with some modifications: briefly, 5 g of sediment was resuspended in 13.5 mL of 100 mM sodium phosphate, 100 mM TrisHCl, 100 mM EDTA pH 8.0, 1.5 M NaCl and 1% CTAB (Extraction buffer) and 1 mg of Protease K. The sediment was incubated at 37 °C for 30 min to allow the Protease K reaction. After, 1.5 mL of SDS 20 % was added at the sediment and it was lysated by Freeze-Thawing (3 cycles) in dry ice and at 65 °C. The lysis was completed by incubation at 65 °C for 1 h. The sample was centrifuged at 6000 x g for 10 min at 20 °C and the supernatant was withdrawn. The pellet was resuspended by adding 4.5 mL of extraction buffer and 0.5 mL of SDS 20% and vortexed for 10 s. After, it was incubated for 10 min at 65 °C and centrifuged as described above. These steps were repeated three times and all supernatants were united. Successively, the total supernatant was resuspended with an equal volume of chloroform:isoamyl alcohol (28:1) and centrifuged for 10 min at 6000 x g. The metagenomic DNA was precipitated in 0.6 x volume of isopropanol for 1 h at room temperature followed by 30 min of centrifugation at 10000 x g. Finally, the DNA was resuspended in water (200 µL) and checked on agarose gel. The amount of DNA was measured at 260 nm on a spectrophotometer (Nanodrop).

Amplification of rRNA 16S genes for DGGE analysis

The archaeal and bacterial rRNA 16S genes were amplified by PCR by using the following degenerate primers:

Bacteria	
341f_GC	5'-CGCCCGCCGCGCCCGCGCCCGI CCCGCCGCCCGCCCG CCTACGGGAGGCAGCAG-3'
907r	5'-CCGTCAATTCMTTGTAGTTT-3'
Archaea	
Arch344f_GC	5'-CGCCCGCCGCGCCCGCGCCCGT CCCGCCGCCCGCCCG ACGGGGYGCAGCAGGCGCGA-3'
Arch915r	5'-GTGCTCCCCGCAATTCCT-3'

Table 3.2.1: Degenerate primers for Archaea and Bacteria. In red GC clamps

The GC clamps were added in the forward primers to enhance the migration on gel electrophoresis. The amplification reaction was performed with the EuroTaq Polymerase (Euroclone) on 150-200 ng of DNA by using the following programme:

3.3- Experimental procedures

Temperature [°C]	Time [min]	Cycles
94	10	1*
94	0.5	2
65	1	
72	1.5	
94	0.5	2
64	1	
72	1.5	
94	0.5	2
63	1	
72	1.5	
94	0.5	2
62	1	
72	1.5	
94	0.5	2
61	1	
72	1.5	
94	0.5	2
60	1	
72	1.5	
94	0.5	2
59	1	
72	1.5	
94	0.5	2
58	1	
72	1.5	
94	0.5	2
57	1	
72	1.5	
94	0.5	2
56	1	
72	1.5	
94	0.5	2
55	1	
72	1.5	

*After this cycle Taq Polymerase was added.

Figure 3.2.1: PCR protocol for rRNA 16S genes amplification.

The PCR products (600 bp) were verified by electrophoresis on 1% agarose gel.

Amplification of rRNA 16S genes for sequencing.

To amplify the metagenomic DNA for rRNA 16S genes sequencing, the universal primers was used: 926wF: 5'-[aaactYaaaKgaattgRcgg]-3' and 1392R: 5'-[acgggcggtgtgtRc]-3'. The amplification reaction was performed with the AccuPrime™ Taq DNA Polymerase High Fidelity (Life technologies) by using the following programme: hot start 5 min at 95 °C; 30 cycles 2 min at 95 °C, 1 min at 56 °C and 1 min at 68 °C; final extension 10 min at 68 °C. The PCR products were verified by electrophoresis on 1% agarose gel.

3.4 – Result and discussion

Recent advances in ‘omics’ technologies, particularly when applied in a systems biology context, have made significant inroads into the prediction of *in situ* functionality of microbial communities (Cowan DA *et al.* 2015). Especially in extreme environments, most microorganisms are recalcitrant to cultivation-based approaches (Amann RI *et al.*, 1995; Lorenz P *et al.*, 2002). Therefore, culture-independent metagenomic strategies are promising approaches to assess the phylogenetic composition and functional potential of microbial communities living in extreme environments (López-López O *et al.*, 2013). In addition, this approach implement tremendously the access to enzymes from (hyper)thermophilic microorganisms that have important potential applications in several biotechnological processes. To these aim in our laboratory metagenomic analysis of hot spring Pisciarelli led us to identify a vast panel of CAZymes involved in the (hemi)cellulose degradation for the second generation bioethanol production. These analyses brought to light the microbial consortia inhabiting Pisciarelli. Currently, Pisciarelli field showed an evident increase of the geothermal activity being affected by a continuous increasing temperature of the existing fumaroles, local seismicity, and occurrence of novel fumaroles mixed with jets of gas and boiling water (Fedele A *et al.*, 2014). In addition, in the Pool along the eastern side of the small hill to the East of this volcanic site, the points of greenhouse gas emissions have increased. In January 2013, partly due to heavy rains, the disappearance of the main fumarole recently opened and the appearance of a vent that emits high-pressure steam and liquid water up to 3-4 meters high were observed (Fedele A *et al.*, 2014). So far it has never been documented if geochemical changes could influence the microbial consortia colonizing these habitats. Therefore, in order to monitor the microbial community variations in Pool1, we undertook an annual campaign of weekly samplings in Pisciarelli to isolate and analyze the metagenomic DNA. In parallel, the geochemical changes were monitored by temperature and pH

measurements and by water composition analysis. In addition, every week we took pictures of Pool1 for monitoring the morphological changes. This multidisciplinary project made use of the skills of several collaborators and the experimental workflow is illustrated in Figure 3.4.1:

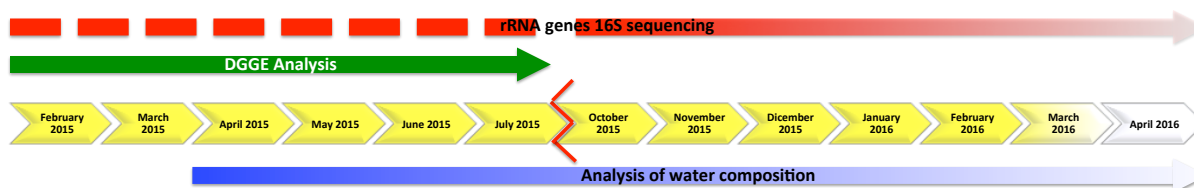



Figure 3.4.1: Experimental workflow. Red dashed arrow: rRNA 16S genes sequencing of some samples from February 2015 to July 2015; red nuanced arrow: rRNA 16S genes sequencing of the sample from July to date. Green arrow: DGGE analysis from February to July; Blue nuanced arrow: analysis of water composition from April 2015 to date. Symbol : Interruption of the samplings.

Geochemical changes monitoring

Samples were collected from February 2015 to date, preleving 2 L of water containing mud from Pool1 (Figure 3.4.2).



Figure 3.4.2. Picture of the sampling.

In situ temperatures ranged from 82 to 94 °C, whereas pH values varied considerably between 2.3 and 6.5. In Figure 3.4.3 the measurements from February 2015 to December 2015 are shown:

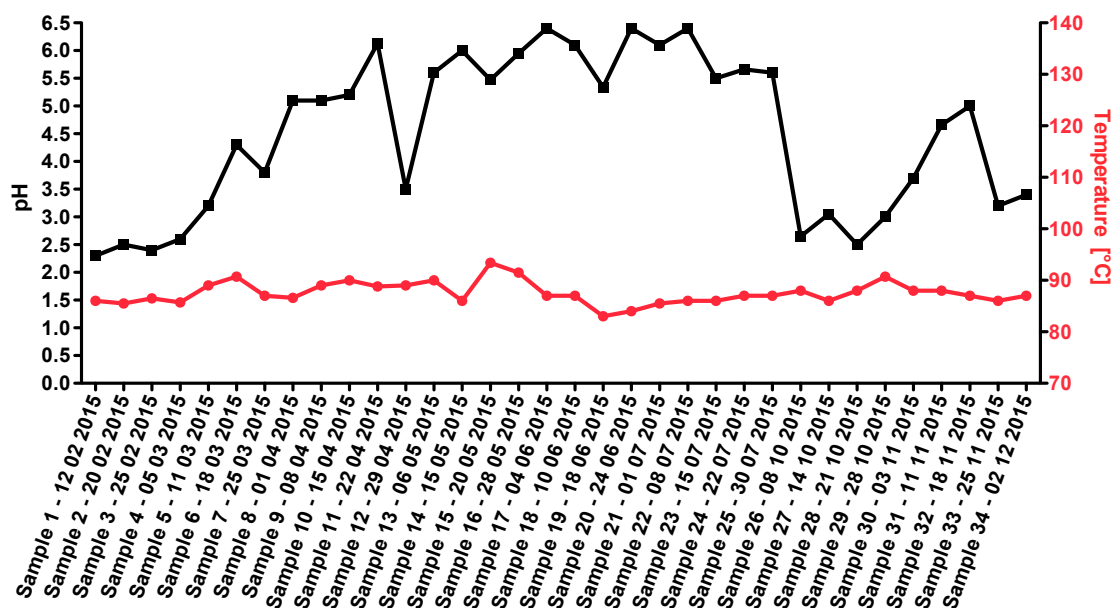


Figure 3.4.3: Temperature and pH variations from February to December 2015.

The major pH variations occurred during rainy weeks when we observed a lowering pH. This phenomenon might be due to the rainwater that reacted with H_2S emission gas by producing SO_2 and H_2SO_4 acidifying Pool1. Moreover, the rainwater might react with the sulfur deposited as crystals on the hill East the Pool1 (Figure 3.4.4).

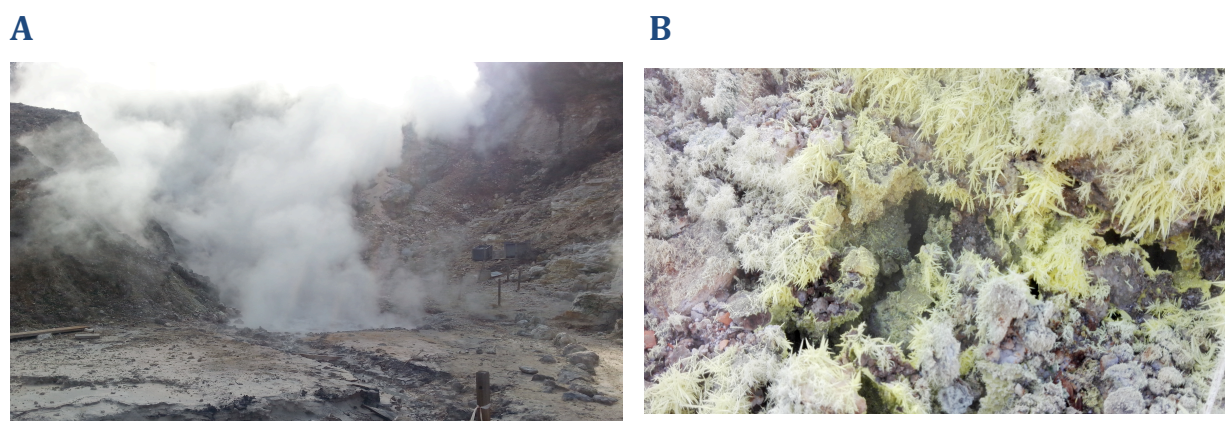


Figure 3.4.4: Pictures of Pisciarelli (A) gas emission near Pool 1; (B) Sulfur crystals

From April 2015 we analyzed the water composition on the filtered sample, in collaboration with Dr Rosario Avino of National Institute of Geophysics and Volcanology (Naples, Italy).

Sample	date	pH titrino	pH	HCO ₃ ⁻ ppm	F ⁻ mg mL ⁻¹	Cl ⁻ mg mL ⁻¹	SO ₄ ²⁻ mg mL ⁻¹	Na ⁺ mg mL ⁻¹	Li ⁺ mg mL ⁻¹	NH ₄ ⁺ mg mL ⁻¹	K ⁺ mg mL ⁻¹	Mg ⁺⁺ mg mL ⁻¹	Ca ⁺⁺ mg mL ⁻¹	δ ¹⁸ O mg mL ⁻¹	δD mg mL ⁻¹
11	22-04-15	5.60	6.13	6.5	0.18	15.73	3463	28.9	n.a.	1154	37.0	0.480	40.0	5.08	12.75
12	29-04-15	3.73	3.50	0.0	0.74	7.21	2583	12.1	n.a.	869	19.5	0.496	47.0	3.94	7.77
13	06-05-15	6.71	5.60	35.5	0.41	4.59	3396	11.9	n.a.	1190	32.5	0.479	48.7	4.43	9.65
14	15-05-15	6.71	6.00	35.7	0.49	7.28	5533	14.2	n.a.	1879	66.8	1.146	98.3	4.36	9.91
15	20-05-15	6.61	5.47	25.9	0.40	4.28	4452	13.2	n.a.	1528	51.4	0.859	83.6	4.21	9.49
16	27-05-15	6.70	5.95	28.5	0.46	5.75	3927	17.0	n.a.	1382	41.1	0.665	59.7	4.43	10.13
17	04-06-15	6.93	6.40	35.4	0.58	3.27	4145	9.4	n.a.	1487	43.6	0.566	61.5	5.14	13.45
18	10-06-15	6.79	6.10	34.5	0.59	3.45	4181	5.5	n.a.	1506	43.5	0.468	58.8	5.40	14.44
19	18-06-15	6.73	5.34	29.5	0.34	16.87	4489	15.3	n.a.	1597	47.7	0.746	47.6	5.44	14.49
20	24-06-15		6.40			3.25	4241	9.7	n.a.	1456	123.6	0.348	49.9	4.80	13.27
21	01-07-15		6.10			3.63	4248	5.4	n.a.	1565	37.7	0.543	49.3	5.03	14.29
22	08-07-15		6.40		0.45	4.75	4241	7.0	n.a.	1362	32.4	0.223	41.2	4.52	10.26
23	15-07-15		5.50		0.47	3.52	4122	11.9	n.a.	1323	30.4	0.168	33.1	4.59	12.32
24	22-07-15		5.66		0.36	2.88	4237	12.3	n.a.	1298	87.9	0.141	33.4	4.75	13.89
25	30-07-15		5.60		0.30	4.15	4995	16.8	n.a.	1530	35.9	0.688	85.9	4.74	13.89
26	08-10-15		2.65		1.44	3.21	3648	8.6	n.a.	987	19.4	0.104	25.4	4.68	13.38

Table 3.4.1: Water composition analysis. In light blue the ions decrease, in dark blue the ions increase.

As shown in Table 3.4.1, Cl⁻ increased considerably in Samples 11 and 19 while NH₄⁺, K⁺, Mg⁺⁺, Ca⁺⁺ increased remarkably in Sample 14 than in Sample 13 while the highest amount of K⁺ was measured in Sample 20. The decrease of NH₄⁺

observed in the Samples 12 and 26 could be causing the lowering of pH in these samples.

Previous studies on the metal composition in Pisciarelli revealed that the main element was iron. In particular, the typical feature of solfataric fields was the existence of two dominating zones in the soil: the upper oxidized zone was often rich in ferric iron. Below the oxidized layer, an anoxic zone exists, often black-colored due to ferrous sulfide. The interphase between these zones is characterized by the presence of elemental sulfur and is therefore slightly yellow-colored. In addition, as in other geothermal pools, arsenate is an important component in Pisciarelli Solfatara (Huber R *et al.*, 2000).

DGGE analysis

To verify the microbial changes from February 2015 to July 2015, we withdrawn 2 L of water from Pool1 and we separated the water from the sediment by centrifugation. From the sediment we extracted the metagenomic DNA (mDNA) by using the protocol reported in Zhou *et al.* 1996, we verified the purity by gel agarose and we measured the amount of mDNA by reading at 260 nm ([Table 3.4.2](#)).

3.4- Results and discussion

Samples	g _· sediment	ng µl ⁻¹	µg _{TOT}	µg _{DNA} g ⁻¹ sediment
1	20	2620	400	20
2	4.6	2300	460	100
3	10	640	128	12.8
4	3.5	542	108	31
5	5	720	144	29
6	4.7	148	30	6
7	4.6	247	49	11
8	6.5	180	36	6
9	2.4	285	57	24
10	6.8	287	57	8
11	6.0	275	55	9
12	6.0	113	22.6	4
13	6.0	214	43	7
14	35.6	ND	ND	-
15	1.1	41.6	4.16	4
16	4	310.9	31.9	8
17	7.1	88	17.6	2
18	5	58.8	11.8	2
19	5	51.3	10.3	2
20	5	165.3	33	7
21	5	302	60.4	12
22	8	180.4	36	5
23	6.5	1170	234	36
24	6.2	191	38.2	6
25	5	509	100	20
26	4.4	816	163	37
27	6.5	474	95	15
28	5.4	4620	974	180
29	5	200	40	8
30	6	977	195	32
31	4.25	216	43.2	10
32	5.9	126.5	25.3	4
33	3.5	282	56.4	16
34	3.4	354.9	71	21

Table 3.4.2: Summary of the amount of mDNA purified from samples withdrawn from February to December 2015.

The mDNA purification was highly variable from minimum of $2 \mu\text{g g}^{-1}_{\text{sediment}}$ up to $180 \mu\text{g g}^{-1}_{\text{sediment}}$. This might be due to sample variable features, indeed the sediment sometimes was composed by higher size particles (Figure 3.4.5) that influenced the sediment weight but from which we extracted a tiny amount of DNA.

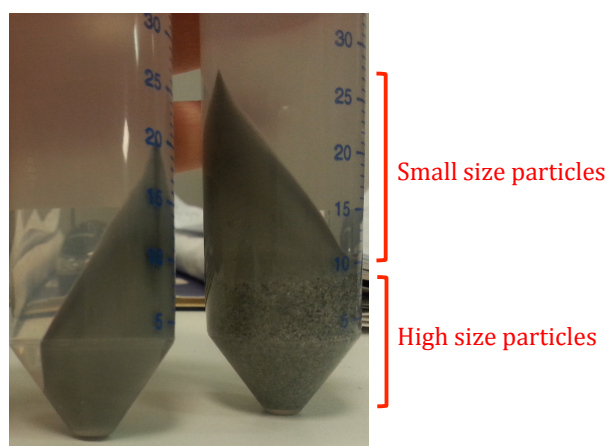


Figure 3.4.5: Sediments obtained after centrifugation of the Samples 1 and 2.

Successively, to analyze the microbial species present in the samples we amplified rRNA 16S genes by using degenerated primers for V4/V6 variable regions amplification for Archaea and Bacteria. The PCR products were checked on agarose gel (Figure 3.4.6). We obtained only the amplicons by using the archaeal primers, as expected by previous metagenomic data where we did not identify any Bacteria within Pool1. Instead, we did not obtain amplification on Sample 14 probably due to low mDNA amount.

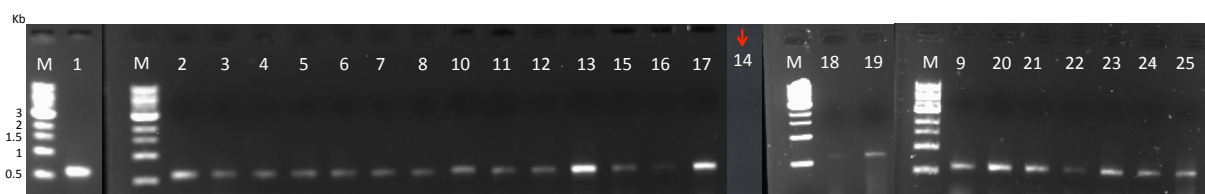


Figure 3.4.6: Agarose gels of PCR products for DGGE. The numbers indicated the sample used as template. M: Marker; B: Blank

The PCR products were analyzed by Denaturing Gradient Gel Electrophoresis (DGGE) in collaboration with group of Prof. Garabed Antranikian of the Institute for Technical Microbiology (TUHH, Hamburg). In DGGE an increasing gradient of concentrations of denaturing chemicals are used to force DNA molecules to unwind. Any variation in DNA sequences will result in different migration, allowing to distinguish the amplicon even by a single nucleotide. The fingerprint achieved by DGGE analysis (Figure 3.4.7) showed low microbial diversity with only one band present in all samples and a second band present in most of them (Samples 3, 8, 9, 10, 11, 12, 13, 15, 16, 17, 23) when Pool1 showed a pH 5.0-6.0. The main band was sequenced and corresponded to *Acidianus* spp, coinciding with our previous metagenomic data.

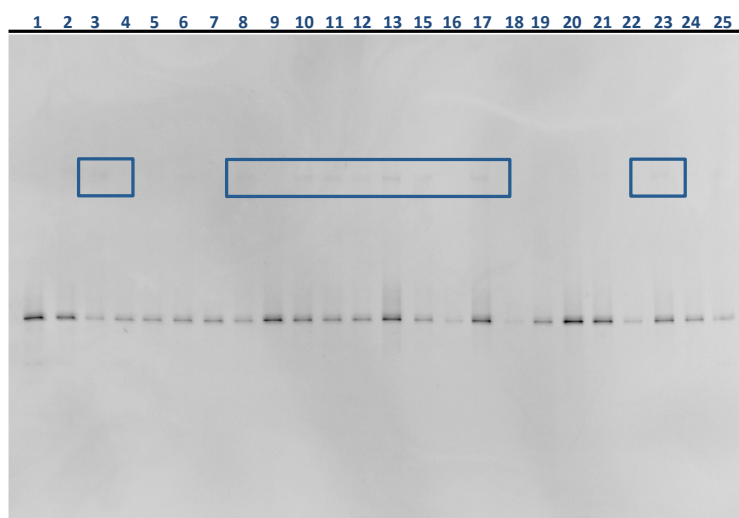


Figure 3.4.7: Fingerprint of the species colonizing Pool1

However, the low sensitivity of this technique, and the uncertainty of the DNA producing the bands, led us to study more in depth the microbial changes thorough rRNA 16S genes sequencing.

rRNA 16S genes sequencing

The sequences of the 16S rRNA genes can provide a detailed comprehensive view of microbial diversity, therefore we embarked in the amplification of mDNA samples with universal primers for archaeal and bacterial rRNA 16S genes. Using universal primers allows to achieve semiquantitative analyses, because the ratio between Archaea and Bacteria is maintained during the PCR. We amplified the metagenomic DNA with degenerated primers (see the Experimental Procedures section) for V6/V8 variable regions of Samples 1, 13, 14, 15, 16, 24 and the amplicons were checked on agarose gel. The Illumina 454 sequencing was performed by MrDNA (www.mrdnalab.com, Shallowater, TX, USA). Operational taxonomic units (OTUs) were classified using BLASTn against a database derived from GreenGenes (<http://greengenes.lbl.gov/cgi-bin/nph-index.cgi>), RDP II (<http://rdp.cme.msu.edu>) and NCBI (www.ncbi.nlm.nih.gov) and were defined by clustering at 3% divergence (97% similarity). The sequencing results were reported in Table 3.4.3.

Identified microorganisms in Pisciarelli						
	PCR.1 12-02-2015	PCR.13 06-05-2015	PCR.14 15-05-2015	PCR.15 20-05-2015	PCR.16 27-05-2015	PCR.24 22-07-2015
<i>sulfolobaceae</i>	305	675	107	1074	901	837
<i>sulfolobus</i>	635	277	63	202	190	282
<i>acidianus ambivalens</i>	6789	4202	836	2518	3027	4042
<i>pyrobaculum</i>	0	554	48	1654	1536	819
<i>thermoproteaceae</i>	0	351	40	908	728	521
<i>pyrobaculum arsenaticum</i>	0	308	38	929	1314	419
<i>metallosphaera spp.</i>	3	3	0	3	0	1
<i>pyrobaculum islandicum</i>	0	10	0	23	14	7
<i>pyrobaculum oguniense</i>	0	1	0	0	4	0
<i>metallosphaera sedula</i>	68	4	0	3	7	0
<i>acidianus</i>	2170	1459	300	1273	1353	1595
<i>desulfurococcaceae</i>	0	19	0	2	6	6

Table 3.4.3: number of genera and species obtained by rRNA 16S genes sequences.

We observed that *Acidianus* was the major genus present in all sequenced samples. In Sample 1 (86 °C and pH 2.3, 12/02/2015) *Acidianus* was followed in abundance by *Sulfolobus*. Sulfolobales maintained constant in all samples, except Sample 14. It is worth noticing that previous papers (Inskeep WP *et al.*, 2010;

Wang S *et al.*, 2014), reported that pH played a key role in selection of species inhabiting high temperature environment. In particular, Inskeep and coworkers in Yellowstone National Park and Wang and collaborators in Tengchong (China) found archaeal species, belonging to the Order Sulfolobales at pH 2-3, while at pH above 6.0 they found bacterial species belonging to the Order Aquificales. Instead, remarkably, in Pisciarelli a pH closer to neutrality did not determinate the presence of Bacteria. In addition, although the pHs were some at the upper limit measured in Pisciarelli (Samples 13, 15 and 24) and others exceeded the pH range at which for *Sulfolobus* grows optimally (0.9-5.8) (Samples 14 and 16) (Brock TD *et al.*, 1972.), we observed the presence of this genus indicating that *Sulfolobus* spp. in Pisciarelli had higher pH tolerance if compared to its homologs elsewhere.

Gas explosion in Pisciarelli

In May 18th 2015 a remarkable gas explosion happened in Pool1, causing morphological and microbiological changes. In particular, during the sampling of May 15th (Sample 14) Pisciarelli we noticed a colour darker than the previous one (Sample 13, May 6th) and that of the following week (Sample 15, May 20th) (Figure 3.4.8). In addition, the mud was highly dense: indeed, while we recovered usually $6-7 \text{ g}_{\text{sediment}} \text{ L}^{-1}_{\text{mud}}$ after centrifugation, in Sample 14 we harvested $50 \text{ g}_{\text{sediment}} \text{ L}^{-1}_{\text{mud}}$. Also the sample collected in May 20th showed a density higher than usual ($35 \text{ g L}^{-1}_{\text{mud}}$), but during the following weeks, density restored at about $6 \text{ g L}^{-1}_{\text{mud}}$.



Figure 3.4.8: Pictures of Pisciarelli before and after gas-explosion

Although the 2 L Sample 14 produced more grams of sediment, only tiny amounts of mDNA was purified. Generally from all the samples the extracted DNA ranged between 2 and 180 $\mu\text{g}_{\text{DNA}} \text{g}^{-1}_{\text{sediment}}$, instead from 35 g of Sample 14 we purified an amount barely detectable on agarose gel ([Figure 3.4.9](#)) and not quantifiable at 260 nm ([Table 3.4.1](#)).

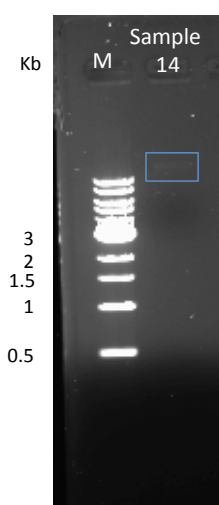


Figure 3.4.9: Agarose gel of DNA extracted by Sample 14

To understand the microbial changes before and after the gas explosion, we sequenced the rRNA 16S genes of Samples 13, 14, 15 and 16 ([Table 3.4.3](#)). In Sample 14 (86 °C pH 6.0, 15/05/2015) we noticed for all species a remarkable decrease in number of sequences, although temperature and pH were concordant with the average of that period ([Figure 3.4.1](#)). In the following week

(Sample 15, 93 °C pH 5.5, 20/05/2015) we observed a noticeable increase of *Pyrobaculum*, becoming the major genus colonizing Pool1 together with *Acidianus*. However, this latter genus recovered completely already in Sample 16 (91 °C pH 6.0, 28/05/2015), dominating Pool1 again. Finally in Sample 24 (87 °C pH 5.7, 22/07/2015), we observed the abundance of the genera in the Order *Acidianus* > *Pyrobaculum* > *Sulfolobus* (Table 3.4.3).

We could speculate that the gas explosion might have gone up on surface *Pyrobaculum*, as it being an anaerobic crenarchaeon generally might grow in deeper zone of Pisciarelli. Indeed, the dark colour showed before the explosion could be due to gone up of anoxic oxidized black-colored layer. Since some species of *Pyrobaculum*, among these *Pyrobaculum arsenaticum*, were not tightly anaerobe, they could grow on the upper layer of Pool1. Indeed, in Sample 15 we found *Pyrobaculum* as dominant genus with *Acidianus*. However, we could not exclude that the drastic microbial changes were due to the chemical conditions in Pisciarelli immediately before the gas explosion. This might explain the drastic reduction of all living forms before the explosion that we observed in Sample 14. For instance, the high accumulations of NH_4^+ and K^+ ions might have inhibit the growth of *Sulfolobus* species by affecting its RNA polymerase activity (Park C B and Lee SB, 1999).

To compare the microbial community variation with the chemical composition of the samples, principal component analysis was performed. The heatmap clearly showed that Sample 14 differed from Samples 13, 15 and 16, clustering separately. Instead, Samples 15 and 16 were similar to each other (Figure 3.4.10).

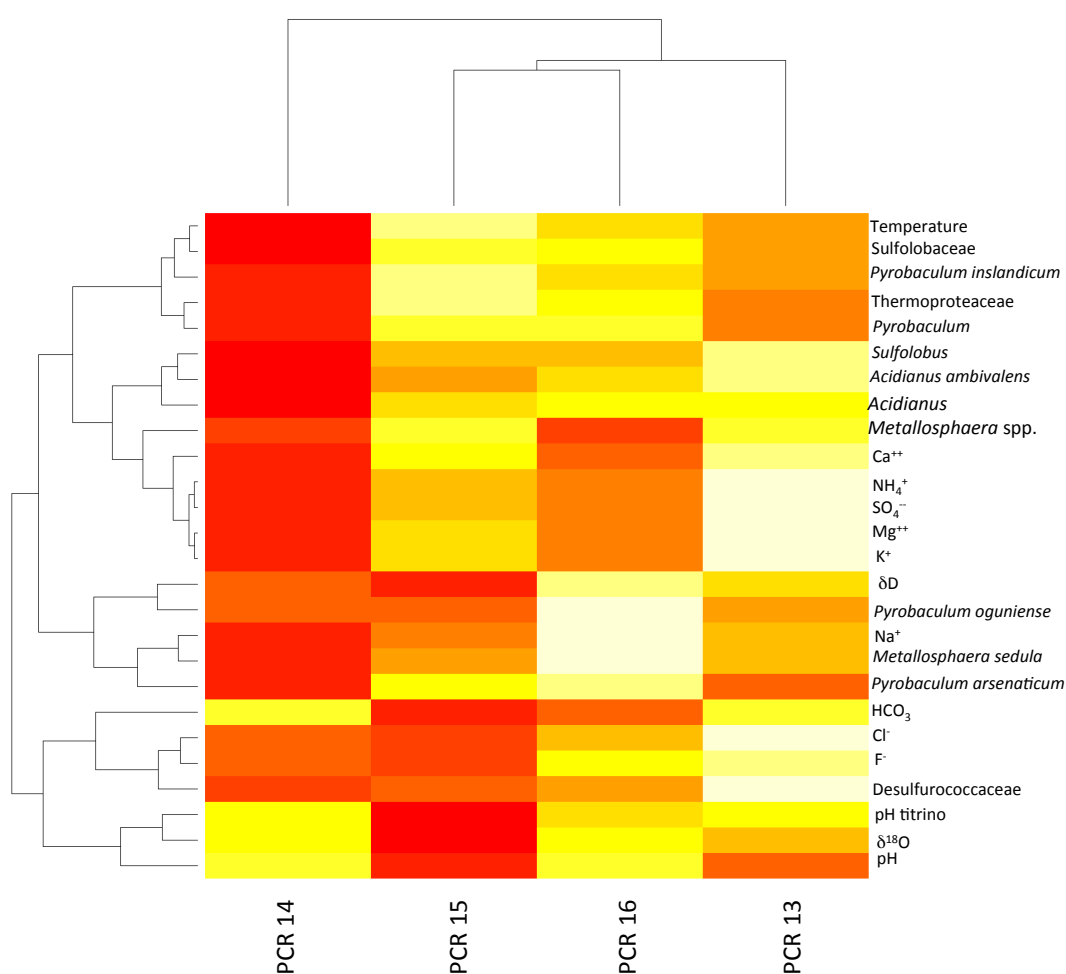


Figure 3.4.10: Heatmap of the Samples 13, 14, 15 and 16.

In conclusion, Pisciarelli Solfatara is an ecological niche colonized by hyperthermophiles with low microbial diversity. This microbiome is exposed to several geochemical variations due to hydrothermal activity of the Solfatara. Indeed, we observed a drastic microbial change caused by a gas explosion that led a remarkable reduction of the microbial population.

Recently, the first correlation between microbial diversity and temporal changes has been analyzed in Tengchong (China) by considering the seasonal changes due to monsoon (Wang S. *et al.*, 2014). Instead, our work represented the first weekly sampling campaign in an extreme environment that monitored the microbial change related to both geochemical and seasonal variations.

3.5- References

- Amann RI, Ludwig W, Schleifer KH. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev.* 1995 Mar;59(1):143-69. Review.
- Briggs BR, Brodie EL, Tom LM, Dong H, Jiang H, Huang Q, Wang S, Hou W, Wu G, Huang L, Hedlund BP, Zhang C, Dijkstra P, Hungate BA. (2013) Seasonal patterns in microbial communities inhabiting the hot springs of Tengchong, Yunnan Province, China. *Environ Microbiol* 16, 1579–1591
- Brock TD, Brock KM, Belly RT, Weiss RL. (1972) *Sulfolobus*: A New Genus of Sulfur-Oxidizing Bacteria Living at Low pH and High Temperature. *Arch Mikrobiol* 84, 54–68.
- Cowan DA, Ramond J-B, Makhalanyane TP, De Maayer P. (2015) Metagenomics of extreme environments, *Current Opinion in Microbiology*, Volume 25, June 2015, Pages 97-102
- Lee CK, Barbier BA, Bottos EM, McDonald IR, Cary SC. (2012) The intervalley soil comparative survey: The ecology of dry valley edaphic microbial communities. *ISME J* 2012, 6:1046-1057.
- Fedele A, Di Giuseppe MG, Troiano A, Somma R, Esposito R, Caputo T, Patella D, Troise C, De Natale G. Geothermal fluids monitoring by time lapse electrical resistivity tomography: the Pisciarelli distribute degassing test site (Campi Flegrei – South Italy) DOI: 10.13140/2.1.4202.1440
- Hobel CFV, Marteinson VT, Hreggvidsson GO, Kristjansson JK. (2005) Investigation of the microbial ecology of intertidal hot springs by using diversity analysis of 16S rRNA and chitinase genes. *Appl. Environ. Microbiol.* 2005, 71, 2771-2776.
- Hou W, Wang S, Dong H, Jiang H, Briggs BR, Peacock JP, Huang Q, Huang L, Wu G, Zhi X, Li W, Dodsworth JA, Hedlund BP, Zhang C, Hartnett HE, Dijkstra P, Hungate BA. A comprehensive census of microbial diversity in hot springs of Tengchong, Yunnan Province China using 16S rRNA gene pyrosequencing
- PLoS One. 2013;8(1):e53350. doi: 10.1371/journal.pone.0053350.
- Huber R, Huber H, Stetter KO. (2000) Towards the ecology of hyperthermophiles: biotopes, new isolation strategies and novel metabolic properties. *FEMS Microbiol Rev.* 2000 Dec;24(5):615-23.

- Inskeep WP, Rusch DB, Jay ZJ, Herrgard MJ, Kozubal MA, Richardson TH, Macur RE, Hamamura N, Jennings Rd, Fouke BW, Reysenbach AL, Roberto F, Young M, Schwartz A, Boyd ES, Badger JH, Mathur EJ, Ortmann AC, Bateson M, Geesey G, Frazier M. (2010) Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function." PLoS ONE 5, no. 3 (2010).
- Isaia R, Marianelli P, Sbrana A (2009), Caldera unrest prior to intense volcanism in Campi Flegrei (Italy) at 4.0 ka B.P.: Implications for caldera dynamics and future eruptive scenarios, *Geophys. Res. Lett.*, 36, L21303, doi:10.1029/2009GL040513.
- López-López O, Cerdán ME and González-Siso MI (2013) Hot Spring Metagenomics *Life* 2013, 2, 308-320; doi:10.3390/life3020308
- Lorenz P, Liebeton K, Niehaus F, Eck J. (2002) Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Current Opinion in Biotechnology*, vol. 13, no. 6, pp. 572–577, 2002.
- Miller, S.R.; Strong, A.L; Jones, K.L; Ungerer, M.C. Bar-Coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *App. Environ. Microbiol.* 2009, 4565 4572.
- Menzel P, Gudbergssdóttir SR, Rike AG, Lin L, Zhang Q, Contursi P, Moracci M, Kristjansson JK, Bolduc B, Gavrilov S, Ravin N, Mardanov A, Bonch-Osmolovskaya E, Young M, Krogh A, Peng X. Comparative Metagenomics of Eight Geographically Remote Terrestrial Hot Springs. *Microb Ecol.* 2015 Aug;70(2):411-24
- Meyer-Dombard DR, Shock EL, Amend JP. (2005) Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. *Geobiology* 2005, 3, 211 227.
- Orsi G, de Vita S, Di Vito M. (1996), The restless, resurgent Campi Flegrei nested caldera (Italy): Constraints on its evolution and configuration, *J. Volcanol. Geotherm. Res.*, 74, 179–214, doi:10.1016/S0377-0273(96)00063-7.

- Park CB, Lee SB. (1999) Inhibitory Effect of Mineral Ion Accumulation on High Density Growth of the Hyperthermophilic Archaeon *Sulfolobus solfataricus*.
J Biotechnol Bioeng 87, 315–319.
- Rosi M, Sbrana A. (1987) Phlegraean Fields, edited by M. Rosi and A. Sbrana, Quad. Ric. Sci., 9(114), 175 pp.
- Wang S, Dong H, Hou W, Jiang H, Huang Q, Briggs BR, Huang L. Greater temporal changes of sediment microbial community than its waterborne counterpart in Tengchong hot springs, Yunnan Province, China. Sci Rep. 2014 Dec 19;4:7479. doi: 10.1038/srep07479.
- Zhang W, Li F, Nie L. (2010) Integrating multiple ‘omics’ analysis for microbial biology: application and methodologies. Microbiology 2010, 156:287-301.
- Zhou J, Bruns MA, and Tiedje JM (1996) DNA Recovery from Soils of Diverse Composition Applied and Environmental Microbiology, Feb. 1996, p. 316–322 0099-2240/96/\$04.00.

Chapter 4 –
Identification and characterization
of the first de-*N*-acetylase from
the hyperthermophilic archaeon
Sulfolobus solfataricus

Chapter 4 of this thesis is focused on the identification of novel carbohydrate active enzymes through functional screening of cellular extracts. The work describes the identification and molecular characterization of the first de-*N*-acetylase from the Crenarchaeon *Sulfolobus solfataricus*. This section begins with a brief introduction on this model organism (paragraph 4.1.1) and the roles played by *N*-acetyl-glucosamine in Sulfolobales (paragraphs 4.1.2). The purpose of this work is shown in paragraph 4.2, while the experimental part is described in paragraph 4.3. In paragraph 4.4 results and discussion are shown. At last, the paragraph 4.5 reports the literature references cited in this chapter.

4.1- Introduction

4.1.1- *Sulfolobus solfataricus*

The thermoacidophilic Archaeon *Sulfolobus solfataricus* (strain P2; DSM1617) is a member of the Order Sulfolobales within the Phylum Crenarchaeota and was isolated from an acid solfataric field (Pisciarelli) near Naples, Italy, more than 35 years ago (Zillig W *et al.*, 1980). In 2001, the genome sequence of *S. solfataricus* has been completed and annotated (She Q *et al.*, 2001). It is the most widely studied organism of the crenarchaeal branch and is a useful model to study DNA replication, cell cycle, chromosomal integration, transcription, RNA processing, and translation (Kort JL *et al.*, 2013) and an interesting source of enzymes of applicative interest (Moracci M *et al.*, 2000; Cobucci-Ponzano B *et al.*, 2008).

S. solfataricus P2 is an aerobic chemoorganoheterotroph with optimal growth at 80°C (range, 60°C to 92°C) and pH 2-4 (Selig M *et al.*, 1997). It degrades glucose via the archaeal-type/modified branched Entner-Doudoroff pathway (Selig M *et al.*, 1997), but it is able to grow heterotrophically on various carbon and energy sources, such as tryptone, peptides and amino acids (Grogan DW, 1989); and many different sugars, including pentoses (e.g., D-arabinose, L-arabinose, and D-xylose), hexoses (e.g., D-galactose, and D-mannose), disaccharides (e.g., maltose and sucrose), and polysaccharides (e.g., starch and cellulose) (Lalithambika S *et al.*, 2012). This is reflected by its noticeably high number of GH genes, such as those involved in the breakdown of starch like the extracellular surface-layer associated α -amylase (SSO1172, GH57) (Worthington P *et al.*, 2003) and the intracellular maltase that releases glucose from maltodextrins (SSO3051, GH31) (Ernst HA *et al.*, 2006; Rolfsmeier M and Blum P, 1995). These sugars are also substrate for the GH13 enzymes TreY (SSO2095) and TreZ (SSO2093) that convert them into trehalose (Maruta K *et al.*, 1996), one of the main compatible solutes found in *Sulfolobus* sp. (Santos H and da Costa MS, 2001). A gene cluster involved in glycogen synthesis and breakdown was also found including a glycogen synthase (SSO0987, GT5), a glycogen associated α -amylase (SSO0988,

GH57) (Cardona S *et al.*, 2001), a α -glucoamylase (SSO0990, GH15) (Kim MS *et al.*, 2004), and a putative glycogen debranching enzyme (SSO0991). Moreover, the genome of *S. solfataricus* encodes three different endoglucanases possibly active on (hemi)cellulose (SSO1354, SSO1949 and SSO2534). The *sso2534* gene product exhibited optimal activity on carboxymethyl cellulose (CMC) (Limauro D *et al.*, 2001) while SSO1354 enzyme showed endoglucanase and xylanase activities (Cannio R *et al.*, 2004; Maurelli L *et al.*, 2008).

The study of *Sulfolobus solfataricus* is of high interest as source of enzymes and proteins with unique properties in terms of thermostability at the operative conditions, but also because this organism displays extensive protein *N*-glycosylation that has been shown to be important in motility, cellular adhesion, cell–cell communication, biofilm formation and maintenance of cellular shape (Jarrell KF *et al.*, 2010; Calo D *et al.*, 2010; Guan Z *et al.*, 2012; Yurist-Doutsch S *et al.*, 2010; Tripepi M *et al.*, 2012). Remarkably, protein *N*-glycosylation is essential, since tunicamycin inhibits cell division by affecting protein glycosylation (Hjort K *et al.*, 2001). The essentiality of this post translational modification might be explained by the stabilizing effect of *N*-glycosylation on the proteins exposed to extreme conditions and demonstrated by the enhanced number of glycosylation sites found in thermophilic glycoproteins relative to their mesophilic counterparts (Meyer BH and Albers SV., 2013). In addition, the abundance of *N*-glycan chains in *Sulfolobus* sp., (Peyfoon E *et al.*, 2010; Palmieri G *et al.*, 2013; Albers SV *et al.*, 2011), suggests that part of the CAZyme repertoire in this organism could be involved in their recycling, similarly to what occurs in Bacteria (Johnson W *et al.*, 2013).

4.1.2 - *N*-acetyl-glucosamine in Sulfolobales

In Sulfolobales, *N*-acetyl-glucosamine is a component of different glycoconjugates such as *N*-glycans decorating proteins and exopolysaccharides. The *N*-glycans in Sulfolobales are unique within the Archaea and show similarities to the eukaryal *N*-linked glycan. Specifically, in these organisms and in another thermoacidophilic archaeon, *Thermoplasma acidophilum*, *N*-glycans are linked to proteins via a *NN*-diacetyl-chitobiose core (GlcNAc- β 1,4-GlcNAc) (Peyfoon E. *et al.*, 2010; Palmieri G. *et al.*, 2013; Yang LL and Haug A, 1979), a common feature of all eukaryal protein *N*-linked glycans (Burda P and Aebi M, 1999). The sugar composition and the stereochemistry of the linkages between the monosaccharides composing Sulfolobales protein-*N*-glycans are still poorly characterized. Recently, the structure *N*-glycan decorating the surface protein SSO1273 from *S. solfataricus*, was partially determined and was shown to be one sugar unit larger than that found in *S. acidocaldarius* (Palmieri G. *et al.*, 2013). The branched heptasaccharide has the structure Hex₄(GlcNAc)₂ plus sulfoquinovose, where the four hexoses are D-mannose and D-glucose similar to the previous structure reported for *S. acidocaldarius*, whose tribranched hexasaccharide is composed of two GlcNAc residues, two Man residues, a sulfoquinovose and a Glc residue (Meyer BH and Albers SV, 2013)(Figure 4.1.1).

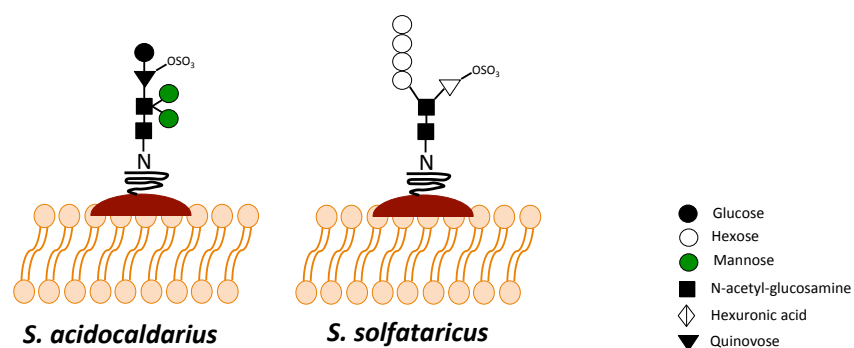


Figure 4.1.1: *N*-glycan in *S. acidocaldarius* and *S. solfataricus*.

Exopolysaccharides (EPSs) make up a substantial component of the extracellular polymers surrounding most microbial cells in extreme environments like Antarctic ecosystems, saline lakes, geothermal springs or deep sea hydrothermal vents. In Bacteria biofilms, *N*-acetylglucosamine has been found in both α - and β -bonds: the pathogenic bacterial strains, including *Staphylococcus epidermidis*, *Staphylococcus aureus*, *E. coli*, *Yersinia pestis*, *Bordetella* sp., *Acinetobacter baumannii*, *Actinobacillus pleuropneumoniae*, *Burkholderia cepacia* complex (Bcc), and *Aggregatibacter actinomycetemcomitans*, produce poly- β -1,6-*N*-acetylglucosamine (Whitfield GB *et al.*, 2015); instead, the pathogenic Gram-negative *Pasteurella multocida* synthesize a group of polymers, called glycosaminoglycan (GAG)-like that are typically less complex than their eukaryotic counterparts due to an absence of modifications such as sulfation (DeAngelis PL *et al.*, 2002). Indeed in Eukaryota, the glucosaminoglycan Heparan sulfate is composed by the repeating disaccharide unit of an uronic acid (glucuronic acids, GlcA, or iduronic acid, IdoA) and *N*-acetylglucosamine GlcA β or IdoA- α 1,4-GlcNAc- α 1, where IdoA and GlcNAc can be modified by sulfation (DeAngelis PL *et al.*, 2013).

EPS biosynthesis is one of the most common protective mechanisms among the adaptation strategies developed by extremophiles, (Nicolaus B *et al.*, 2010). Indeed, in Sulfolobales, EPSs enhance the adaptation to rapid variations in temperature, pH and geochemical conditions. The amount of EPSs synthesized by *Sulfolobus* sp. vary based on environmental pH and temperature: *S. acidocaldarius* increases by 80% EPSs synthesis when it grows at pH 6.0 while 5-fold increase biofilm formation is detected at growth temperatures lower than 60°C. Moreover, this Archaeon uses EPSs for cell adhesion on surfaces and as a carbohydrate reservoir to exploit as energy source or during biofilm formation (Koerdt A *et al.*, 2010). The *S. solfataricus* EPS contain glucose, mannose, galactose and *N*-acetyl-glucosamine, but the regio- and stereochemistry of the glycosidic bonds between these monosaccharides is still unknown (Zolghadr B

et al., 2010; Nicolaus B *et al.*, 2003). In this context, the identification of enzymes that catalyze the modification of *N*-acetylglucosamine in *Sulfolobales* might shed light on the biosynthesis and degradation of this important glycoside and its derivatives.

4.2- Purpose of work

The metabolism and the biological function of *N*-glycans decorating proteins and EPS in Archaea is still poorly understood and one of the keys to access information on these important biomolecules is through the identification and functional characterization of CAZymes involved in their modification. Recently, in our laboratory, we identified in *Sulfolobus solfataricus* P2 a novel glycoside hydrolase belonging to family GH116, which is involved in the hydrolysis of β -*N*-acetyl-glucosaminides and its detailed characterization increased our knowledge on the phylogeny of this class of enzymes (Ferrara MC *et al.*, 2014).

This part of my PhD project was aimed to identify novel CAZymes from the same source, involved in the modification of α -*N*-acetyl-glucosamine. To date no GH with α -*N*-acetylglucosamidase activity have been identified in Archaea; thus, the first part of my work was focused on the isolation and identification of an enzymatic activity able to hydrolyze the chromogenic substrate 4Np- α -GlcNAc from *S. solfataricus* P2. In the second part, the biochemical characterization of the purified enzyme and on its recombinant form allowed to unveil a novel and unexpected carbohydrate esterase.

4.3- Experimental Procedures

Archaea and Bacterial Strain

Sulfolobus solfataricus P2 (She et al., 2001);

E. coli strains used in this work are TOP10: F- *mcrA* Δ (*mrr-hsdRMS-mcrBC*) Φ 80*lacZ* Δ M15 Δ *lacX74 recA1 araD139* Δ (*araleu*)7697 *galU galK rpsL* (StrR) *endA1 nupG* and BL21 star: (DE3) .F⁻ *ompT hsdSB* (rB⁻ mB⁻) *gal dcm rne131*(DE3) Invitrogen.

Culture media

S. solfataricus Brock medium (Brock TD *et al.*, 1972) adjusted to pH 3.5 with sulfuric acid and supplemented with yeast extract, sucrose and casaminoacids (0.1 % each) as carbon source.

LB (Luria-Bertani Broth) (1 liter): 10 g NaCl, 5 g yeast extract, 10 g tryptone

Reagents

All commercially available substrates (2Np- and 4Np- α -GlcNAc, 4Np- β -GlcNAc, UDP-GlcNAc, GlcNAc-1P, GlcNAc-6S, GalNAc, ManNAc, *NN*-diacetyl-chitobiose) were purchased from Sigma-Aldrich and Carbosynth unless otherwise designated.

Standard growth conditions

S. solfataricus P2 was grown at 80 °C, pH 3.5 in Brock's salt medium supplemented with yeast extract, sucrose, and casamino acids (0.1% each) (Brock TD *et al.*, 1972). The growth of cells was monitored spectrophotometrically at 600 nm and the cells were harvested at the early stationary phase (1.0 optical density) by centrifugation at 5000×g for 20 min at 4°C.

Native Protein Purification

A culture of 9.0 L of *S. solfataricus* P2 was centrifuged and the cellular pellet (9 gr) was resuspended in 2 mL g⁻¹ cells of Tris-HCl 50 mM pH 8.0 buffer supplemented with 0.1% Triton X-100. The cells were incubated with Lysozyme and Benzonase for 60 min at 37 °C and then lysated mechanically with five cycles of French Press and centrifuged at 10,000 ×g for 30 min at 4 °C. The free cell extract (FCE) was loaded on a High Load 16/10 Q-Sepharose High Performance column (GE-Healthcare) equilibrated in 50 mM Tris-HCl buffer, pH 8.0 (Buffer A) at a flow rate of 2 mL min⁻¹. The run was performed with an initial step of extensive wash with Buffer A (2-column volumes) followed by a linear ionic strength gradient from 0 to 1 M NaCl in Buffer A (3-column volumes) and a final step with 1 M NaCl in Buffer A (2-column volumes). At these conditions, the de-*N*-acetylase activity was found primarily in the fractions eluted at about 300 mM NaCl. Active fractions were pooled and dialyzed versus 20 mM phosphate buffer pH 7.0. The dialyzed pool was equilibrated in 1 M ammonium sulphate and loaded on a HiLoad 26/10 Phenyl Sepharose High performance (GE-Healthcare), equilibrated at a flow rate of 3 mL min⁻¹ with phosphate buffer pH 7.0 supplemented with 1 M ammonium sulphate. After 1 column volume of loading buffer, the protein was eluted with a two-step gradient of phosphate buffer pH 7.0 (0–80%, 2-column volumes; 80–100%, 3-column volumes) followed by a final step at 100% of phosphate buffer pH 7.0 (2-column volumes); the protein eluted in about 90% phosphate buffer pH 7.0. Active fractions were pooled, dialyzed against PBS buffer (20 mM sodium phosphate buffer, pH 7.3, 150 mM NaCl), and concentrated by ultrafiltration on an Ultracon 10K (cut off 10,000 Da). After concentration, the sample was loaded on a Superdex 75 HR 10/300 gel filtration column (GE-Healthcare) and the run was performed at a flow rate of 0.5 mL min⁻¹ in PBS buffer. Active fractions were pooled and concentrated. The de-*N*-acetylase activity was followed performing the enzymatic assay on the chromogenic substrate 4Np- α -GlcNAc at 65 °C for 16

hours. The enzymatic activity was detected taking advantage of the absorbance of 4Np- α -GlcN at 420 nm.

After this procedure SSO2901 was more than 95% pure by SDS-PAGE, stained with SYPRO Orange.

Assay of native SSO2901

The native SSO2901 was assayed in phosphate buffer 50 mM pH 6.5 for 16 hours at 65°C on different substrates 4Np- α -GlcNAc, 4Np- β -GlcNAc and GlcNAc, all at the concentration of 5 mM. The analysis of the reaction products was performed by High Performance Anion Exchange chromatography with Pulsed Amperometric Detection (HPAEC-PAD) equipped with a PA200 column (Dionex, USA). HPAEC-PAD analyses were performed at a flow rate of 0.5 mL min⁻¹ in isocratic elution of 10 mM NaOH for 20 min.

Cloning of recombinant rSSO2900 and rSSO2900

The sso2900 and sso2901 genes were amplified by PCR from the genome of *S. solfataricus* P2 (180 ng) using the synthetic oligonucleotides TOPOFWD2900 (5'-CACCATGGATATTTTAGCAGTAGTT-3') and TOPOREV2900His (5'-CTTTTTTGCAAAGGCTTTTATATCTT-3') TOPOFWD2901 (5'-CACCATGCAAAAAAGTAGTTGTAT-3') and TOPOREV2901His (5'-ACCTTTAGTCCAAAAACCTCGT-3'), for sso2900 and sso2901 respectively. The amplification reaction was performed with the PfuUltra II Fusion HS DNA Polymerase (Stratagene) by using the following programme: hot start 5 min at 95 °C; 10 cycles 1 min at 95 °C, 1 min at 50 °C and 1 min at 72 °C; 30 cycles 1 min at 95 °C, 1 min at 55°C, and 1 min at 72°C; final extension 10 min at 72 °C.

The PCR products obtained were verified by electrophoresis on 1% agarose gel, purified by PCR Kleen spin columns (Biorad) and then cloned in the expression vector pET101/D-TOPO (Invitrogen), in which the respective ORFs are under

the control of the isopropyl-1-thio- β -D-galactopyranoside (IPTG) inducible T7 RNA polymerase promoter and the C-terminal of the proteins were fused to V5 epitope and a 6xHis tag. TOPO cloning was performed following the instruction of the manufactures; briefly, 80 ng (sso2901) and 4 ng (sso2900) of fragments were incubated in salt solution with 15-20 ng vector (1 μ l) at room temperature for 30 minutes. Reactions were used to transform One Shot TOP10 chemically competent cells according to the protocol of the manufacturer (see below for more details). Positive clones were selected through PCR colony and the absence of mutations was verified by sequencing.

The recombinant plasmids pET101/D-TOPO-SSO2901 and pET101/D-TOPO-SSO2901, were transformed into One Shot® TOP10 chemically competent *E. coli* BL21 (DE3) Star provided by the kit. After addition of 20 ng of each plasmid into a vial, the cells were incubated on ice for 30 minutes. Then, they were subjected to a heat-shock at 42°C for 30 seconds and immediately transferred to ice and supplemented with 250 μ L SOC medium. The cells were incubate at 37°C for 1 hour; then 200 μ L of culture were spread on LB-agar plates containing 50 μ g mL⁻¹ ampicillin and incubated for 16 hours at 37 °C.

E. coli cells transformed were used to prepare 100 mL culture. The cells were incubated at 37 °C and when reached an A_{600nm} of 0.6 optical densities the culture was divided in different aliquots for each condition (Figure 4.3.1):

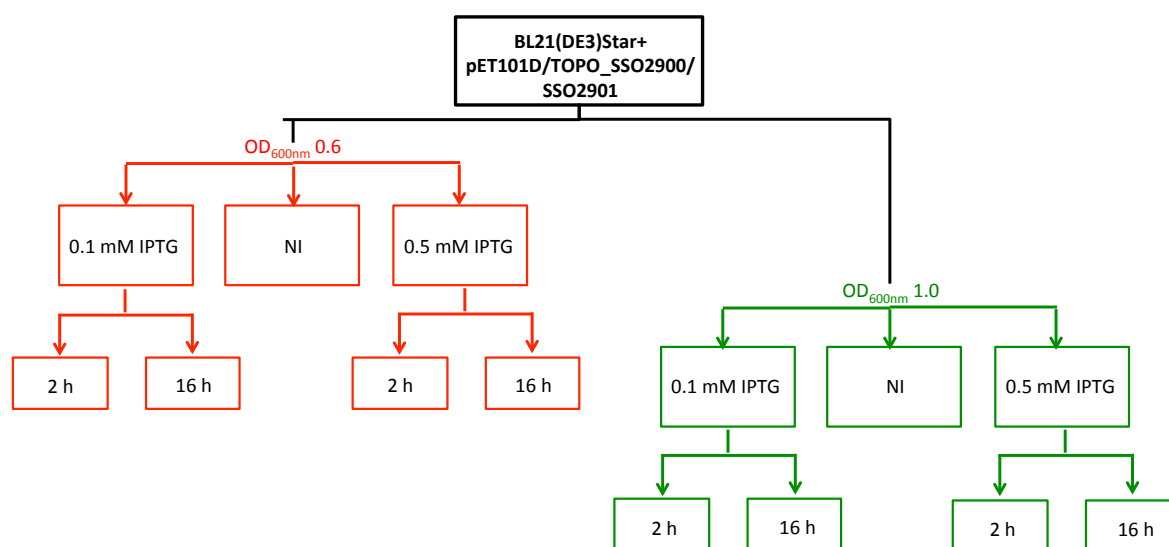


Figure 4.3.1: Diagram of expression trials; NI: cells not induced.

The cultures were harvested by centrifugation at 5,000 × g for 2 minutes and resuspended in loading buffer 1X. Then, cells were loaded on SDS-PAGE (12%) to evaluate the protein expression.

Purification of rSSO2900 and rSSO2901

E. coli BL21 star (DE3) cells were grown at 37 °C in 0.5 L of Luria–Bertani (LB) broth supplemented with ampicillin (50 µg mL⁻¹). The induction of rSSO2900 recombinant expression was performed by adding 0.1 mM IPTG when the growth was at 1 OD, while rSSO2901 was induced by adding 0.5 mM IPTG when the growth has reached 0.6 OD.

Growths for rSSO2900 and rSSO2901 expression were allowed to proceed for 16 h and 2 h, respectively. The cells were harvested by centrifugation at 5000 ×g. The resulting pellets were resuspended in 50 mM sodium phosphate buffer pH 8.0, 300 mM NaCl with a ratio of 5 mL g⁻¹ cells and then were incubated at 37 °C for 1 h with 20 mg of lysozyme (Fluka) and 25 U g⁻¹ cell of Benzonase (Novagen). The cells were lysed by French Press treatment and cell debris were removed by centrifugation at 10,000 × g for 30 min. The FCEs were loaded on

Protino® Ni-IDA Packed Columns (Macherey-Nagel) according to the manufacturer. The eluted fractions of rSSO2900 were pooled and dialyzed against PBS 1X pH 7.3 while the rSSO2901 fractions against 25 mM HEPES-NaOH pH 7.5.

Molecular mass determination of rSSO2901

The molecular mass of recombinant SSO2901 was determined by gel filtration on a Superdex 75 HR 10/300 FPLC column (GE-Healthcare). Molecular weight markers were BSA (66 kDa), ribonuclease A (13.7 kDa), aprotinin (6.5 kDa) and vitamin B12 (1.3 kDa).

Fluorescamine-based de-N-acetylase activity assay

Purified rSSO2901 was tested for de-N-acetylase activity using Fluorescamine fluorophore. Assay mixture containing 50 mM HEPES pH 8.5 and 20 mM GlcNAc was pre-equilibrated at 70 °C prior to the addition of rSSO2901 (3 µg). After 5 minutes, the reaction mixture (150 µl) was quenched by addition of 20% trichloroacetic acid (TCA, 50 µl), which precipitates the protein. After centrifugation (16000 g, 10 min), 125 µl of supernatant were diluted with 375 µl borate 1 M (pH 9.0), and reacted with 10 mM FSA (150 µl in CH₃CN, Invitrogen). After 10 min, the resulting fluorescence was measured (excitation 395 nm, emission 495 nm) using a Fluorescence Spectrometer JASCO FP-8600 (Jasco). The observed increase in fluorescence (fluorescence units [FU] min⁻¹) was converted into mmol/min using a glucosamine (GlcN) standard curve (50-150 mM) (Huang X and Hernick M, 2011). One enzymatic unit is defined as the amount of enzyme catalysing the conversion of 1 µmole of substrate into product in 1 min, at standard conditions.

For determination of the steady-state parameters, activity was measured at 18 different concentrations of GlcNAc (0–80 mM), and the parameters k_{cat} , K_M , and

k_{cat}/K_M were obtained by fitting the Michaelis–Menten equation to the initial linear velocities using the curve-fitting program GraphPad Prism.

Temperature, pH and metals dependance

The temperature profile of rSSO2901 activity was determined in the range of 40–85 °C on 20 mM GlcNAc in buffer HEPES 50 mM pH 8.5 for 5 min. Thermal stability was evaluated by incubating pure rSSO2901 in HEPES buffer, pH 7.5, at different temperatures (60°C, 70°C, 80°C). At intervals, aliquots of 3 µL (3 µg enzyme) were recovered, transferred in ice and assayed at the conditions described above. The residual activity was expressed as a percentage of the maximal enzymatic activity measured before the incubation at indicated temperatures; pH optimum was determined by assaying rSSO2901 in 50 mM of different buffers (Citrate/Phosphate; Phosphate; HEPES; Borate; CHES) at different pHs (5.0-10.0) on 20 mM GlcNAc at 70 °C for 5 min. The metal dependence was verified for Zn^{2+} , Mg^{2+} , Mn^{2+} , Co^{2+} and EDTA. The enzyme was incubated for 5 min at 70 °C with 1 mM of each metal. Then, the standard assay was performed. Similarly, the effect of metal ions on the activity of rSSO2901 was tested by adding the compounds indicated directly in the reaction mixture.

Substrate specificity

The assays of rSSO2901 on different substrates (GlcNAc 20 mM, GalNAc 20 mM, ManNAc 20 mM, GlcNAc-1P 5 mM, GlcNAc-6S 5 mM, 4Np- α -GlcNAc 5 mM, 4Np- β -GlcNAc 5 mM, UDP-GlcNAc 20 mM) were performed by using 3 µg of enzyme at 70 °C in buffer HEPES 50 mM pH 8.5 for 5 min. The analysis of the reaction products was performed by fluorescamine assay and HPAEC-PAD, as described above.

rSSO2900 was assayed on 10 mM UDP-GlcNAc in 50 mM Tris-HCl pH 7.5 with 30 µg of enzyme for 16 hours at 65 °C. HPAEC-PAD analyses were performed at a

flow rate of 0.5 mL min^{-1} in isocratic elution of 10 mM NaOH for 5 min followed by two further isocratic steps (15 mM NaOH for 5 min and 70mM NaOH for 40 min).

4.4- Results and discussion

Glycosidases play a key role in a broad range of biological processes and industrial applications. In particular, the enzymes isolated from hyperthermophilic microorganisms are interesting models to understand the molecular basis of protein stability and, in Archaea, an instrument to define the metabolism of the *N*-glycans and EPS. To this aim, in our laboratory, searches in *S. solfataricus* P2 genome sequence led to the discovery of a novel GH38 α -mannosidase and a GH116 bi-functional β -glucosidase/xylosidase possibly involved in protein *N*-glycosylation and EPS turnover (Cobucci-Ponzano B *et al.*, 2010a; Cobucci-Ponzano B *et al.*, 2010b). Part of my PhD project was aimed to identify novel extremophilic CAZymes through functional screening of cellular extracts of this Archaeon. By following this approach we recently identified and characterized a novel GH116 exo- β -GlcNAcase. This was a remarkable result, considering that no enzymatic activities of this kind were known before among the rich repertoire of *S. solfataricus* CAzymes shown in Figure 4.4.1 (Ferrara MC *et al.*, 2014).

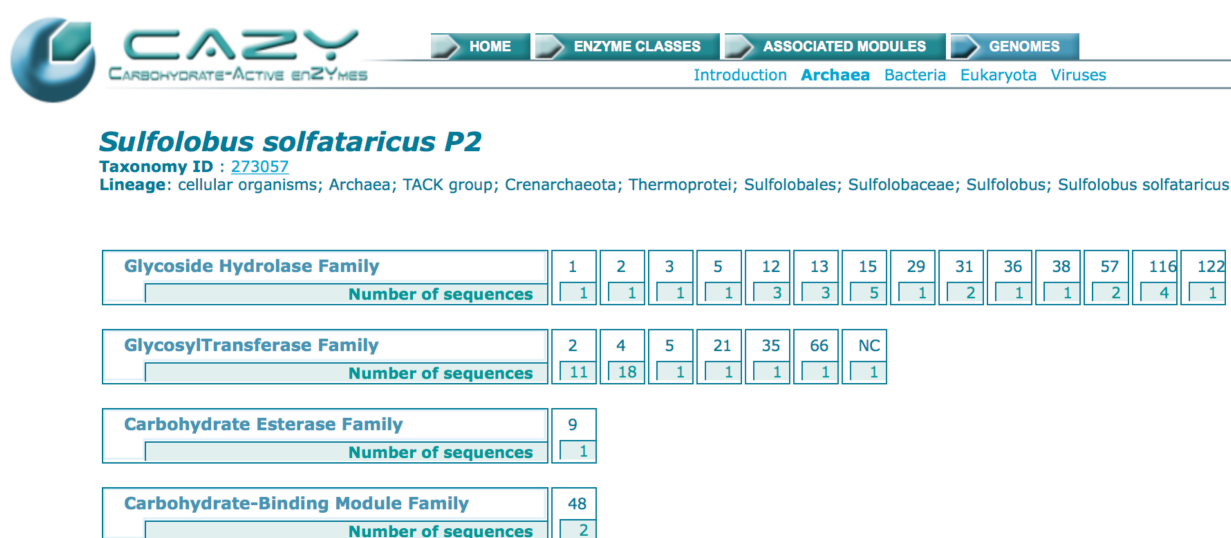


Figure 4.4.1: Numbers of CAZymes present in *S. solfataricus* P2.

In this scenario, we embarked in the search of an α -*N*-acetyl-glucosaminidase

activity from *S. solfataricus* P2, assaying its cellular lysate on 4Np- α -GlcNAc. This enzymatic activity (EC 3.2.1.50), which is classified only in GH89 family, is unknown in Archaea.

Free cell extracts (FCE) were prepared from a cell culture stopped in the early stationary phase. Assays at high temperatures on 4Np- α -GlcNAc allowed to detect tiny amounts of *N*-acetylglucosaminidase activity. Consequently, our efforts were focused on the isolation and identification of this putative *N*-acetyl- α -glucosaminidase. Thus, a procedure was set up enabling the purification of the enzyme of interest through three chromatographic steps (anionic exchange, hydrophobic, and gel filtration chromatographies). Although the purification was performed from 9 L of growth, this procedure allowed to get only tiny amounts of pure protein, detectable by SDS-PAGE only through SYPRO Orange staining. (Figure 4.4.2).

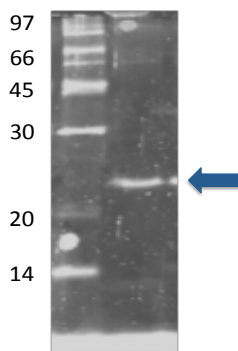


Figure 4.4.2: SDS-PAGE stained with Sypro Orange; lane 1: Marker, lane 2: gel-filtration fraction

Mass spectrometry analysis

To identify the ORF encoding for *N*-acetyl- α -glucosaminidase activity we used a mass spectrometry approach. This study was performed in collaboration with Prof. P. Pucci and Dr. A. Flagiello of the University of Naples “Federico II”. Mass spectrometry analysis was performed on the band excised from the gel loaded

with the sample obtained from the final step of purification. The protein band was hydrolyzed in solution by trypsin and the peptide mixture obtained was fractionated by HPLC and analyzed by tandem mass spectrometry (MS/MS). The MS/MS spectra submitted to MASCOT search led to scores that allowed the clear identification of the hypothetical protein, codified by ORF sso2901 (Table 4.4.1).

identified ORF	Matches	Score	Functional annotation
sso2901	19	392	hypothetical protein
sso2523	2	33	Long-chain-fatty-acid--CoA ligase (fadD-2)
sso2072	1	33	hypothetical protein
sso0879	2	20	hypothetical protein
sso1736	2	16	hypothetical protein

Table 4.4.1: list of protein identified with LC-MSMS

The hypothetical protein SSO2901 showed a MW (22342 Da) matching the migration of the band on gel. In order to understand its putative activity we performed an alignment of the amminoacidic sequence using the program BLASTp. According to this alignment, SSO2901 encoded for a putative de-*N*-acetylase. It is worth noticing that none of the proteins identified by LC-MSMS showed similarity to putative *N*-acetyl- α -glucosaminidases. Therefore, either the identified SSO2901 showed both de-*N*-acetylase and *N*-acetyl- α -glucosaminidase activities or the observed hydrolysis of 4Np- α -GlcNAc occurred through a different mechanism. To verify these hypotheses we performed enzymatic assays of the purified SSO2901 on 4Np- α -GlcNAc, GlcNAc and 4Np- β -GlcNAc and analyzed the reaction mixtures by HPAEC-PAD (Figure 4.4.3)

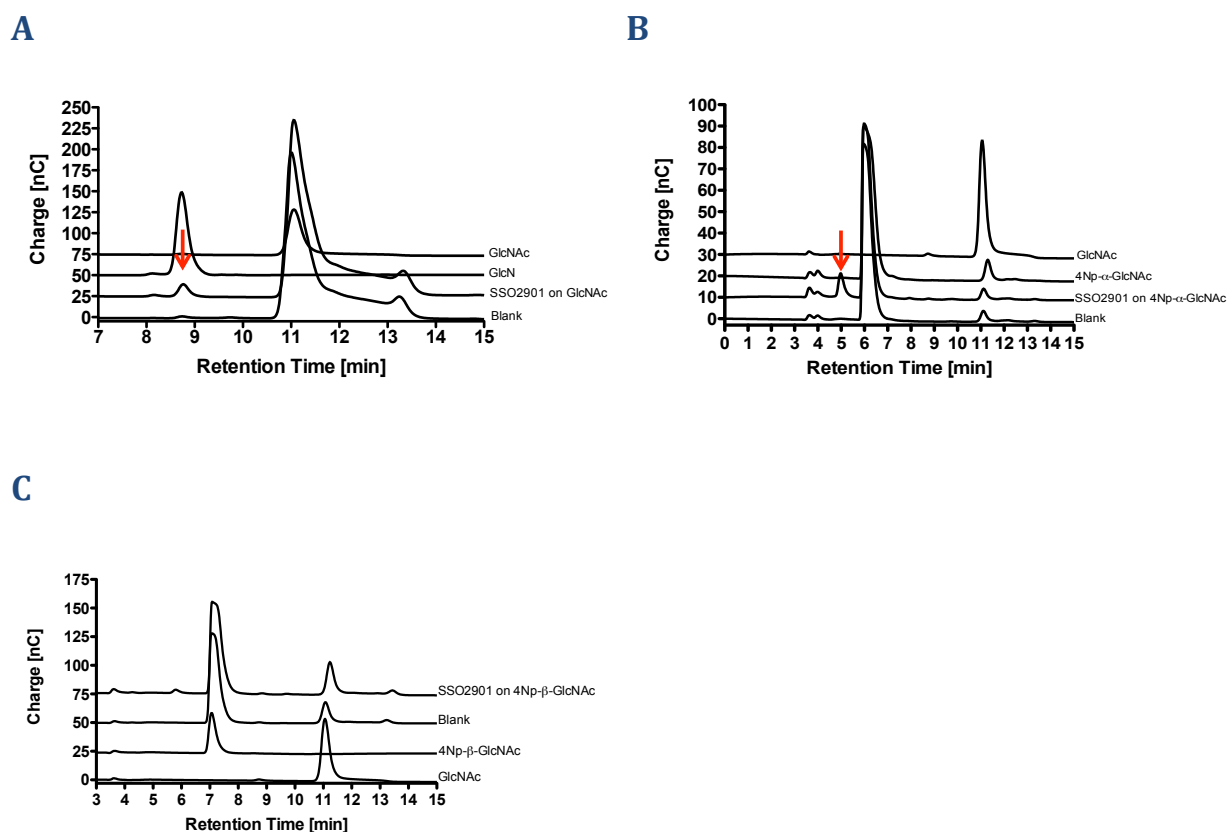


Figure 4.4.3: HPAEC-PAD of products of native SSO2901 on (A) GlcNAc. (B) 4Np- α -GlcNAc and (C) 4Np- β -GlcNAc

The reaction on GlcNAc confirmed that SSO2901 is a de-*N*-acetylase, observing the glucosamine production, while the reaction on 4Np- α -GlcNAc showed the formation of a peak probably corresponding to 4Np- α -GlcN. Instead, remarkably, SSO2901 was not able to hydrolyze the 4Np- β -GlcNAc as the peak corresponding to GlcNAc in the reaction mixture containing the enzyme was present also in the blank and could be due to the spontaneous hydrolysis of the substrate in the conditions of the assay. The explanation of the observed increase in absorbance at 420 nm upon incubation of SSO2901 and 4Np- α -GlcNAc, was that 4Np- α -GlcN possessed at high temperature spectrophotometrical properties similar to 4-nitrophenol (4Np-OH), absorbing at the same wavelength. To confirm that, we performed a chemical de-*N*-

acetylation of both 2Np- and 4Np- α -GlcNAc by incubating these compounds in 100 mM NaOH for 10 min at RT and at 70°C. Surprisingly, only the reactions performed at 70°C absorbed at 420 nm. We analysed the reaction mixtures by HPAEC-PAD and we confirmed the formation of the de-*N*-acetylated products. These results, together with the de-*N*-acetylation of 2Np- α -GlcNAc used as control, are displayed in Figure 4.4.4.

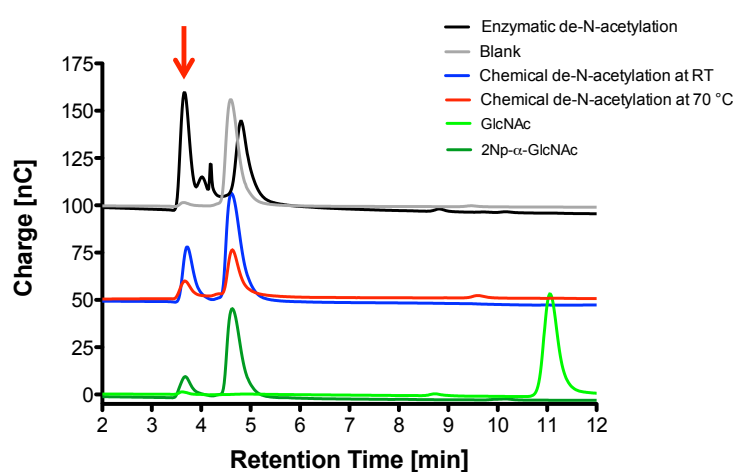


Figure 4.4.4: HPAEC-PAD of chemical de-*N*-acetylation of 2Np- α -GlcNAc

Hence, surprisingly, the spectrophotometric properties of 2Np- and 4Np- α -GlcN allowed us to purify and identify the de-*N*-acetylase SS02901 by following its activity on 2Np- and 4Np- α -GlcNAc.

α -*N*-acetyl-glucosaminidase activity in *S. solfataricus*

In order to understand if the absence of an α -GlcNAcase among the identified proteins by LC-MSMS was due to the absence of this activity in *S. solfataricus* or if the SS02901 activity on 4Np- α -GlcNAc did not allow to purify an eventual α -GlcNAcase, we tested the *S. solfataricus* extracts on 4MU- α -GlcNAc, as the de-*N*-acetylase did not produce a fluorescent product on this substrate. *S. solfataricus* P2 hydrolyzed the 4MU- α -GlcNAc (Figure 4.4.5), suggesting its ability to

metabolize α -N-acetyl-glucosaminides.

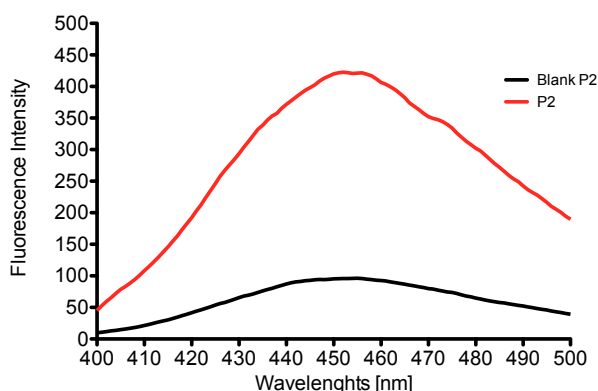


Figure 4.4.5: Fluorescence spectra of *S. solfataricus* P2 extract assayed on MU- α -GlcNAc

Successively, we investigated if this hydrolysis was due to an auxiliar activity of other enzymes, to a cooperative activity between two or more enzymes or an α -GlcNAcase activity. To verify the first hypothesis we tested the α -glucosidases of *S. solfataricus* that, catalyzing the hydrolysis of α -(1,4)-glucosides, might be active also on GlcNAc derivatives. Both the α -glucosidases known, belonging to family GH31, have been characterized, MalA codified by SS03051 (Rolfmeier M. *et al.*, 1998) and XylS codified by SS03022 (Moracci M. *et al.*, 2000). However, these enzymes did not show activity on 4MU- α -GlcNAc (Figure 4.4.6). Therefore, we verified the second hypothesis by assaying the concerted activity of SS02901 with MalA or XylS, since the latter enzymes could be active on α -glucosaminides produced by SS02901. MalA showed activity on 4MU- α -GlcNAc in presence of SS02901, suggesting that *S. solfataricus* used their concerted action to hydrolyze the α -N-acetyl-glucosaminides (Figure 4.4.6).

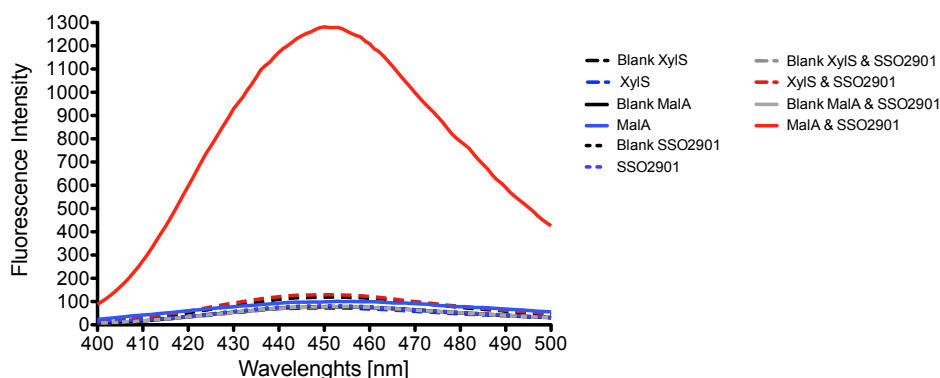


Figure 4.4.6: Fluorescence spectra of reactions on MU- α -GlcNAc

This results indicated that it could be very difficult to verify the presence of an α -GlcNAcase in *S. solfataricus*, even if we could not exclude this eventuality. Moreover, the ability of this microorganism to hydrolyze α -*N*-acetyl-glucosaminides led us to speculate the hypothetical presence of this sugar in glycoconjugates.

It is worth mentioning that the concerted action of SSO2901 and MalA to degradate α -*N*-acetyl-glucosaminides was similar to a mechanism adopted by hyperthermophilic archaeon *Thermococcus kodakarensis* to degradate the chitin. Indeed, in the last steps of the chitinolytic pathway, the de-*N*-acetylase (Tk-Dac) removed the *N*-acetyl group from *NN*-diacetyl-chitobiose (GlcNAc- β -1,4-GlcNAc) producing the disaccharide GlcNAc- β -1,4-GlcN, that was hydrolyzed by β -D-glucosaminidase Tk-GlmA (Tanaka T *et al.*, 2004).

Sequence analysis of SSO2901

According to BLASTp alignments, SSO2901 encoded for a putative GlcNAc de-*N*-acetylase, belonging to LmbE-like superfamily. This superfamily comprised a

series of metal-dependent deacetylases that use a single metal ion to catalyze the hydrolysis of various substrates with a *N*-acetylglucosamine moiety (Viars S *et al.*, 2014). LmbE-like superfamily included *N*-acetylglucosaminyl-phosphatidylinositol (GlcNAc-PI) de-*N*-acetylases (EC 3.5.1.89) from mammals (Nakamura N *et al.*, 1997), yeast (Watanabe R *et al.*, 1999), and protozoa (Chang T *et al.*, 2002) involved in the biosynthesis of GPI anchor; 1-D-myo-inosityl-2-acetamido-2-deoxy- α -D-glucopyranoside (GlcNAc- Ins) de-*N*-acetylase from *Mycobacterium tuberculosis* (MshB) (EC 3.5.1.89) (Newton GL *et al.*, 2000), involved in mycothiol biosynthesis; *N,N*-diacetylchitobiose de-*N*-acetylase from archaeon *Thermococcus kodakaraensis* KOD1 (Tk-Dac) (EC 3.5.1.-) involved in the chitin degradation (Tanaka T *et al.*, 2004), and an antibiotic de-*N*-acetylase from *Actinoplanes teichomyceticus* (Orf2), involved in the modification of lipoglycopeptidic antibiotic teicoplanin (Zou Y *et al.*, 2008).

The GlcNAc-Ins and *N,N*-diacetylchitobiose de-*N*-acetylases are classified into family CE14 of Carbohydrate Esterase in CAZy database. To date, only archaeal and bacterial sequences belong to this family. Instead, antibiotic de-*N*-acetylases are not yet classified in CAZy (Lombard V *et al.*, 2014).

We performed a multialignment with SSO2901 and all CE14 members by using ClustalW software and we constructed a phylogenetic tree based on their sequence similarity (Table 4.4.2 and Figure 4.4.7).

4.4- Results and discussion

GenBank accession	Carbohydrate esterase of family CE14	% of identity
SS02901	Sulfolobus solfataricus P2 SS02901	100
gi 478753334 gb AGJ61906.1	LmbE family protein [Sulfolobus islandicus LAL14/1]	77
gi 323473889 gb ADX84495.1	LmbE family protein [Sulfolobus islandicus REY15A]	77
gi 323477087 gb ADX82325.1	LmbE family protein [Sulfolobus islandicus HVE10/4]	76
gi 227455944 gb ACP34631.1	LmbE family protein [Sulfolobus islandicus L.S.2.15]	76
gi 227458590 gb ACP37276.1	LmbE family protein [Sulfolobus islandicus M.14.25]	76
gi 238379984 gb ACR41072.1	LmbE family protein [Sulfolobus islandicus M.16.4]	76
gi 329566899 gb AEB95004.1	LmbE family protein [Metallosphaera cuprina Ar-4]	42
gi 676315815 gb AIM27447.1	LmbE family protein [Metallosphaera sedula]	39
gi 145702320 gb ABP95462.1	LmbE family protein [Metallosphaera sedula DSM 5348]	39
gi 258591057 emb CBE67352.1	protein of unknown function [Candidatus Methylophilus oxyfera]	26
gi 589090768 gb AHK47521.1	hypothetical protein OV14_b1365 [Ensifer adhaerens OV14]	26
gi 674294741 gb AIL09640.1	glcNac-PI de-N-acetylase family protein [Stenotrophomonas maltophilia]	26
gi 532221016 gb AGT91049.1	hypothetical protein O5Y_05890 [Rhodococcus erythropolis CCM2595]	25
gi 257478761 gb ACV59080.1	LmbE family protein [Alicyclobacillus acidocaldarius subsp. acidocaldarius DSM 446]	25
gi 20516944 gb AAM25112.1	conserved hypothetical protein [Caldanaerobacter subterraneus]	25
gi 610526855 dbj BAO62259.1	PI deacetylase family protein [Pseudomonas protegens Cab57]	25
gi 269789698 gb ACZ41839.1	LmbE family protein [Thermobaculum terrenum ATCC BAA-798]	25
gi 338804072 gb AEJ00314.1	LmbE family protein [Nitrosomonas sp. Is79A3]	24
gi 500242538 gb AGL84696.1	hypothetical protein PFLCHA0_c29260 [Pseudomonas protegens CHA0]	24
gi 646234415 gb AIB36775.1	acetylglucosaminylphosphatidylinositol deacetylase [Pseudomonas simiae]	24
gi 15621886 dbj BAB65879.1	putative diacetylchitobiose deacetylase [Sulfolobus tokodaii str. 7]	24
gi 471210638 gb AGI07388.1	Hypothetical Protein XCAW_01588 [Xanthomonas citri subsp. citri Aw12879]	24
gi 219544070 gb ACL25808.1	LmbE family protein [Chloroflexus aggregans DSM 9485]	24
gi 402255147 gb AFQ45422.1	bacillithiol biosynthesis deacetylase BshB2 [Desulfosporosinus meridiei DSM 13257]	24
gi 384069438 emb CCH02648.1	LmbE family protein [Fibrella aestuarina BUZ 2]	24
gi 83756842 gb ABC44955.1	hypothetical conserved protein [Salinibacter ruber DSM 13855]	24
gi 570730368 gb AHF01588.1	GlcNac-PI de-N-acetylase [Thioalkalimicrobium aerophilum AL3]	24
gi 549146032 emb CDF61751.1	putative PI- deacetylase [Xanthomonas fuscans subsp. fuscans]	24
gi 345445708 gb AEN90725.1	PI de-N-acetylase enzyme, LmbE family [Bacillus megaterium WSH-002]	23
gi 427352230 gb AFY34954.1	LmbE family protein [Calothrix sp. PCC 7507]	23
gi 163667240 gb ABY33606.1	LmbE family protein [Chloroflexus aurantiacus J-10-fl]	23
gi 222447555 gb ACM51821.1	LmbE family protein [Chloroflexus sp. Y-400-fl]	23
gi 193089328 gb ACF14603.1	LmbE family protein [Chloroherpeton thalassium ATCC 35110]	23
gi 146347945 gb EDK34481.1	Conserved hypothetical protein [Clostridium kluyveri DSM 555]	23
gi 402252142 gb AFQ42417.1	putative LmbE-like protein [Desulfosporosinus meridiei DSM 13257]	23
gi 299539852 gb ADJ28169.1	LmbE family protein [Nitrosococcus watsonii C-113]	23
gi 194309306 gb ACF44006.1	LmbE family protein [Pelodictyon phaeoclathratiforme BU-1]	23
gi 5458963 emb CAB50449.1	Hypothetical protein PAB1341 [Pyrococcus abyssi GE5]	23
gi 359351054 gb AEV28828.1	putative LmbE-like protein [Sphaerochaeta pleomorpha str. Grapes]	23
gi 269789613 gb ACZ41754.1	LmbE family protein [Thermobaculum terrenum ATCC BAA-798]	23
gi 46196769 gb AAS81185.1	hypothetical cytosolic protein [Thermus thermophilus HB27]	23
gi 55772582 dbj BAD71023.1	conserved hypothetical protein [Thermus thermophilus HB8]	23
gi 383509325 gb AFH38757.1	bacillithiol biosynthesis deacetylase BshB1 [Thermus thermophilus J-L-18]	23
gi 430010085 gb AGA32837.1	group 2 family protein [Thioalkalivibrio nitratireducens DSM 14787]	23
gi 663511528 gb AIF02121.1	LmbE family protein [uncultured marine group II/III euryarchaeote KM3_153_G11]	23
gi 294801162 gb ADF38228.1	conserved hypothetical protein [Bacillus megaterium DSM 319]	23
gi 719643445 gb AIW83859.1	bacillithiol biosynthesis deacetylase BshB2 [Bacillus weihenstephanensis]	23
gi 163863642 gb ABY44701.1	LmbE family protein [Bacillus weihenstephanensis KBAB4]	23
gi 94551957 gb ABF41881.1	LmbE-like protein [Candidatus Koribacter versatilis Ellin345]	23

Table 4.4.2: % identity of SS02901 against the archaeal CE14 calculated with ClustalW

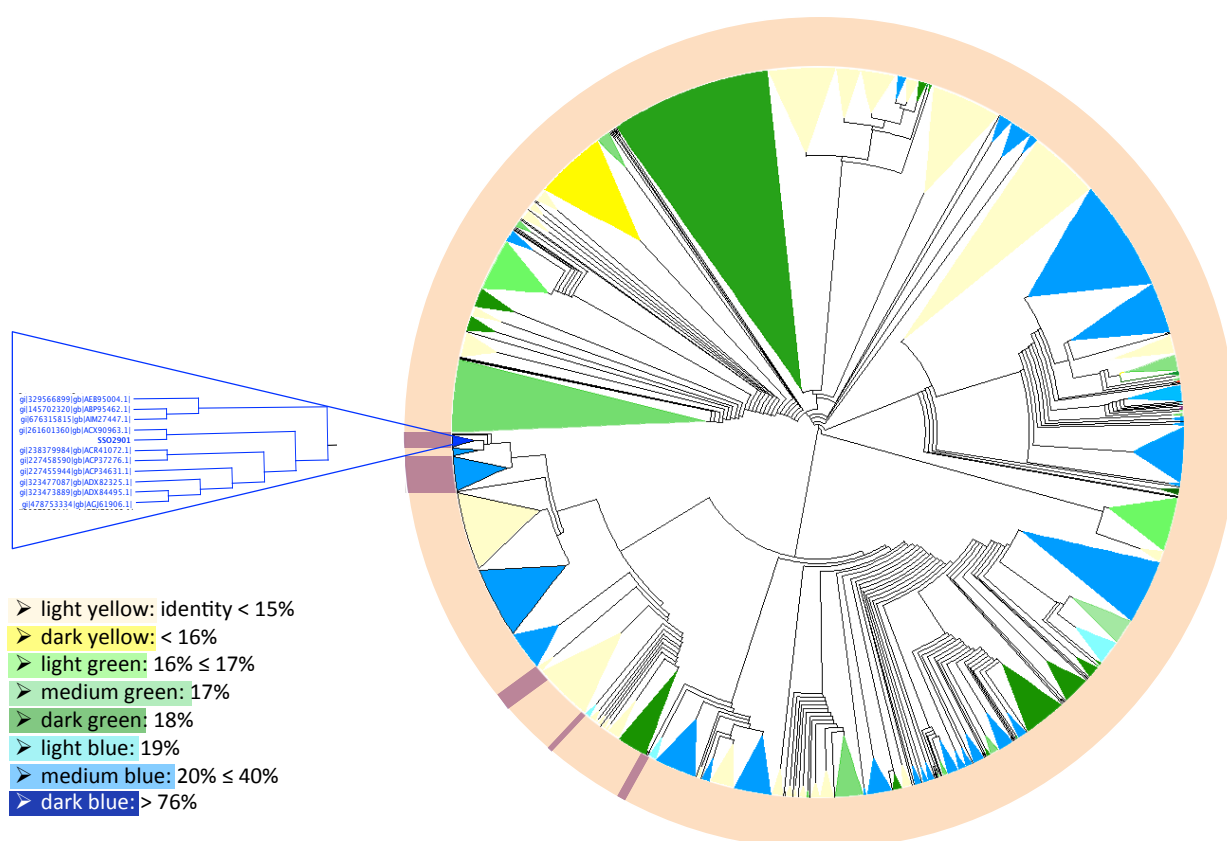


Figure 4.4.7: Phylogenetic tree of SSO2901 and CE14 family. The sequences were clustered based on % of identity with SSO2901, indicated with the colours. The orange outline indicates the bacterial sequences. The purple indicates the archaeal.

Both the multialignment and the phylogenetic tree showed that SSO2901 shared a high sequence similarity with the members of CE14. Nevertheless, surprisingly, the *S. solfataricus* enzyme has not been classified into this family in CAZy database so far. Possibly because CE classification in CAZy is less frequently updated because of the substrate promiscuity of esterases toward non-carbohydrate substrates.

Only eight members of CE14 were characterized. In particular, the archaeal *N,N*-diacetylchitobiose deacetylases from *Haloferax mediterranei*, *Pyrococcus furiosus*, *Pyrococcus horikoshii*, and *Thermococcus kodakarensis* and, the de-*N*-acetylases from the bacteria *Bacillus cereus*, *Mycobacterium tuberculosis* and

Mycobacterium smegmatis. The archaeal de-*N*-acetylases were involved in the degradation pathway of chitin a polysaccharide composed of (β -1,4)-*N*-acetylglucosamine (Khor E, 2001). To our knowledge, *S. solfataricus* is not able to grow on chitin and no hypothetical chitinases were found in its genome. By contrast, bacterial CE14 played a key role in mycothiol and bacillithiol biosynthesis. These low-molecular weight thiols served as analogous of glutathione, functioning as both the primary reducing agent of these organisms and a cofactor for the detoxification of xenobiotics (Newton GL *et al.*, 2008; Hernick M, 2013). Instead, *S. solfataricus* P2 produces glutathione for the maintenance of redox homeostasis (Pedone E *et al.*, 2006). Therefore, the putative metabolic pathway involving SS02901 is completely unknown. To understand the hypothetic role of SS02901 *in vivo*, we analysed the genomic enviroment of sso2901 gene (Figure 4.4.8).

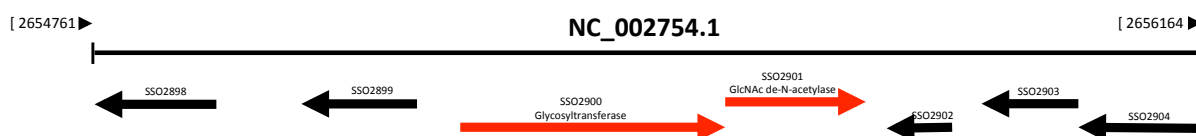


Figure 4.4.8: Genomic enviroment of sso2901.

From the genome, we identified the ORF sso2900 immediately upstream of sso2901 and in the same reading frame. sso2900 encoded for a putative glycosyl transferase belonging to family GT4. A large number of different activities were grouped in this family, such as sucrose synthase, threalose synthase and UDP-GlcNAc transferase. Therefore, we considered it appropriate to characterize sso2900 in an effort of clarifying the role *in vivo* of SS02901 and provide information on GT4 family.

Cloning, expression, and purification of rSS02900 and rSS02901

To shed light on the biological role of SSO2900 and SSO2901 we produced and characterized the recombinant form of both enzymes. To this aim, *sso2900* and *sso2901* genes were cloned in the expression vector pET101/D-TOP0 obtaining the recombinant constructs pET101/D-TOP0-SSO2900 and pET101/D-TOP0-SSO2901, which allowed to express the recombinant enzymes rSSO2900 and rSSO2901 fused to the V5 epitope and to a 6xHis tag at the C-terminal ends. To determine which conditions allowed to obtain the best yields for both enzymes, expression trials were performed in *E. coli* BL21 star (DE3). SDS-PAGE analysis showed that the highest expression of rSSO2900 could be obtained by inducing for 16 hours with 0.1 mM IPTG when the growth has reached 1 OD. Instead, the best yields of rSSO2901 were achieved after induction for 2 hours with 0.5 mM IPTG when the growth has reached 0.6 OD.

The purification procedures consisted of IMAC chromatography step and led to an enzyme preparation >95% pure of both enzymes (Figure 4.4.9). The final yields corresponded to 6 mg per liter of *E. coli* for rSSO2901, while for rSSO2900 we obtained 2 mg per liter of growth (Figure 4.4.9).

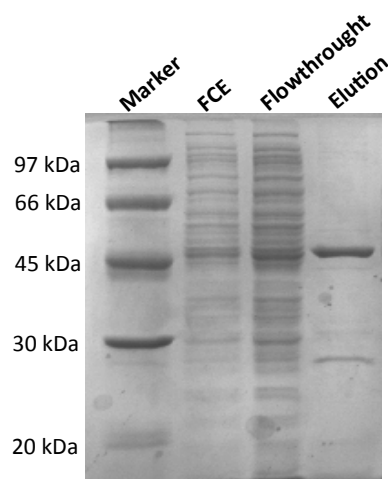
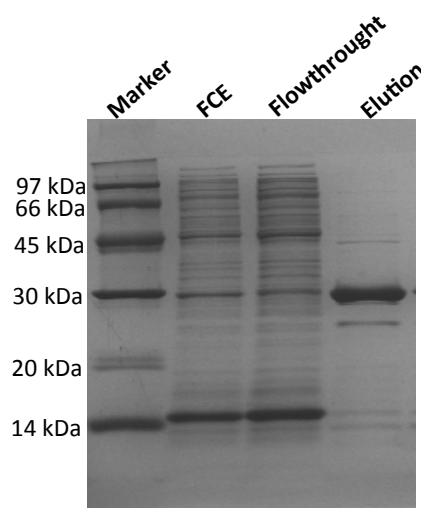
A**B**

Figure 4.4.9: SDS-PAGE of rSSO2900 (**A**) and rSSO2901 (**B**) purifications.

Fractions containing rSS2900 were dialysed against PBS 1X pH 7.3 while pure rSSO2901 was dialysed against 25 mM HEPES pH 7.5.

Characterization of rSSO2901

The molecular weight of the recombinant enzyme in solution was of 21.4 kDa \pm 1.48 kDa comparable to that observed for the denatured monomer. This indicated that rSSO2901 was a monomer in solution.

The biochemical analyses were aimed to reveal the optimal conditions for the catalysis, such as pH and temperature optima, and thermal stability. To determine the pH optimum of rSSO2901, the enzyme was assayed in 50 mM of different buffers in the range of pH 5.0-10.0 on 20 mM GlcNAc. The yields of glucosamine obtained by de-acetylation was measured by using Fluorescamine assay (FSA), as described in the Experimental Procedures paragraph. As illustrated in [Figure 4.4.10](#), rSSO2901 was optimally active in buffer HEPES at pH 8.5.

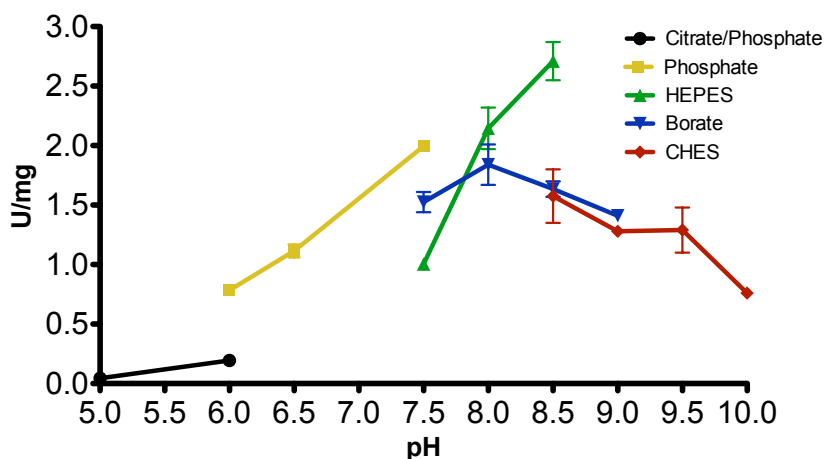


Figure 4.4.10: pH dependence of rSSO2901

All LmbE-like superfamily members possessed the catalytic metal Zn^{2+} in the active site, pentacoordinated by the conserved His-X-Asp-Asp motif. During the

catalysis the intermediate was stabilized by the positively charged Zn and by the imidazolium side chain of His (Broadley SG *et al.*, 2012). To evaluate if rSSO2901 possessed a catalytic metal, we performed the enzymatic assay with EDTA. In addition, since the CE14 deacetylases BC1534 and BC3461 from *Bacillus cereus* were inhibited by 1 mM Zn^{2+} (Deli A. *et al.*, 2010), we tested the enzymatic activity in the presence of metals (Figure 4.4.11). The assays were performed in two different ways: (i) each metal was added directly to the reaction mixture (filled bars), otherwise the enzyme was incubated with 1 mM concentration of different metals for 5 min at 70°C. After the incubation, the enzymatic activity was analyzed in 50 mM HEPES pH 8.5 on 20 mM GlcNAc (pattern bars). rSSO2901 lost about 50% of its activity when incubated with EDTA, suggesting that a catalytic metal was chelated thereby reducing the activity, as shown in other members of LmbE-like superfamily (Viars S *et al.*, 2014). In addition, rSSO2901 was inhibited by Mn^{2+} and Zn^{2+} , as previously reported for the rat phosphatidylinositol (PI)-GlcNAc de-*N*-acetylase (Punta M *et al.*, 2012) and BC1534 and BC3461 from *B. cereus* (Deli A. *et al.*, 2010)

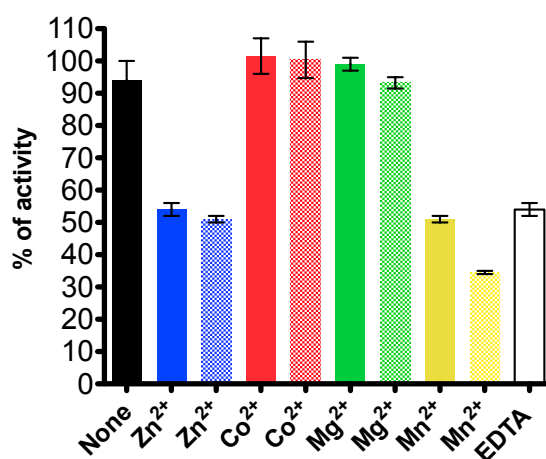


Figure 4.4.11: Enzymatic assay of rSSO2901 in presence of different metals. Filled bar: the metal is added to the reaction mix. Bar with pattern: the enzyme was pre-incubated with metal before the reaction.

The temperature profile of the de-*N*-acetylase activity was assessed in the range 40-85°C, on 20 mM GlcNAc in 50 mM buffer HEPES pH 8.5. As expected for an enzyme from a thermophilic microorganism, the thermophilic curve indicated the maximal activity at 70°C (Figure 3.4.12)

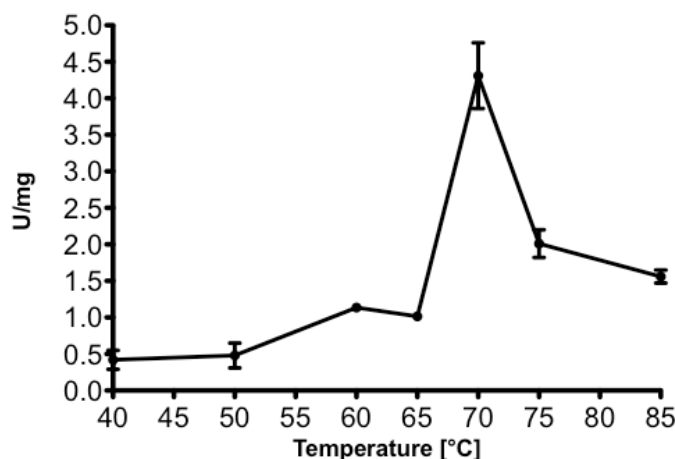


Figure 4.4.12: Thermophilic curve of rSSO2901

In addition, rSSO2901 showed a remarkable stability when incubated for 100 min at 60 and 70°C, maintaining the 50 % of its activity; instead, the enzyme was completely inactivated after 5 minutes at 80 °C (Figure 4.4.13).

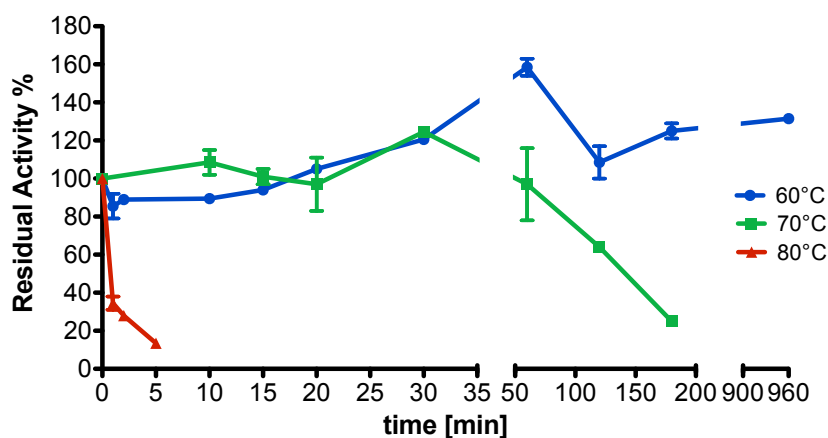


Figure 4.4.13: Thermostability of rSSO2901

On the basis of this initial characterization, the standard assay used in all the following study was performed on 20 mM GlcNAc, in 50 mM HEPES buffer pH 8.5 at 70°C.

Substrate Specificity of rSSO2901

To determine the substrate specificity of rSSO2901, the enzyme was assayed on various *N*-acetyl-glycosides, such as 4Np- α - and 4Np- β -GlcNAc, GlcNAc-1P, GlcNAc-6S, GalNAc, ManNAc, UDP-GlcNAc and *NN*-diacetyl-chitobiose. Enzymatic assays on these substrates indicated that rSSO2901 was active only on GlcNAc, differently to the bacterial de-*N*-acetylases from *B. cereus* and *B. anthracis*. Indeed, these enzymes showed a broad substrate specificity, resulting active on GlcNAc and the oligosaccharides derived therein (chitobiose > chitotriose > chitotetraose) as well as on GlcNAc-1P, GlcNAc-6P, GalNAc, and ManNAc (Deli A. *et al.*, 2010). Moreover, although the archaeal de-*N*-acetylases were mainly active on *NN*-diacetyl-chitobiose (Tanaka T *et al.*, 2004; Mine S *et al.*, 2014) rSSO2901 did not display activity on this substrate even after 16 hours incubation (Figure 4.4.14C). The assays on aryl-*N*-acetyl-glycosides were carried out in 50 mM phosphate pH 6.5 at 65°C for 16 h, the same conditions used for the native enzyme (Figure 4.4.14A and B). The recombinant form was active on 4Np- α - and β -GlcNAc; instead, it is worth mentioning that native SSO2901 did not show activity on 4Np- β -GlcNAc. This could be due to the amount of enzyme used in assay; indeed, although the native enzyme was purified from 9 L of culture, the amounts obtained were too low to be quantified. However, the activity of rSSO2901 on 4Np- α - and 4Np- β -GlcNAc indicated that both anomers were tolerated in the active site, in agreement with the de-*N*-acetylase from *Trypanosoma brucei* that was active on both GlcNAc- α - and GlcNAc- β -PI (Smith TK *et al.*, 2001). In addition, this could suggest that the lack of activity on *NN*-diacetyl-chitobiose could be not due to the β - conformation of the glycosidic

bond in this substrate, but possibly to the fact that the active site did not recognize GlcNAc dimers.

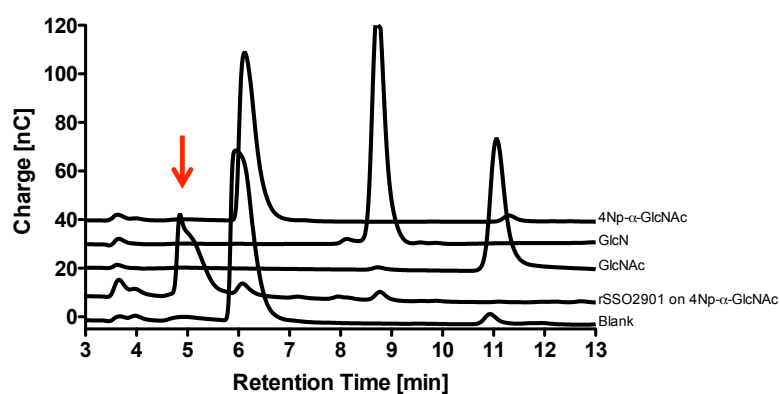
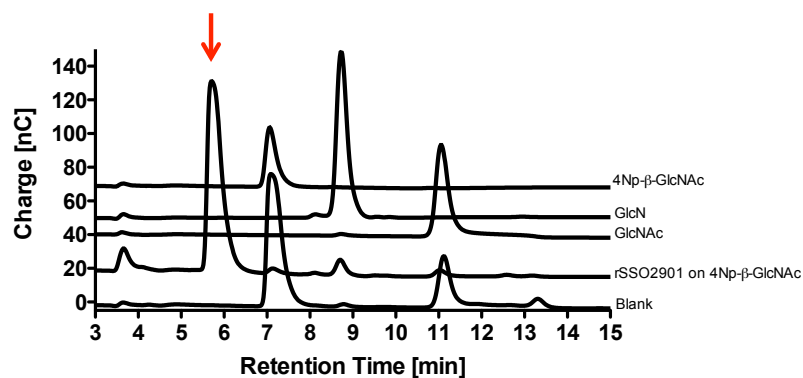
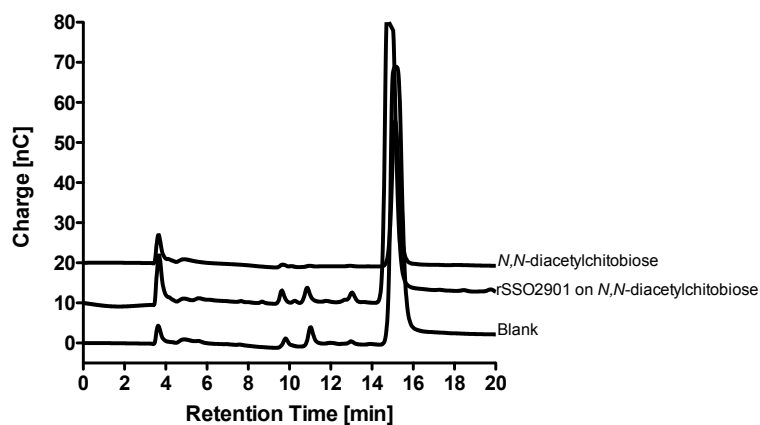
A**B****C**

Figure 4.4.14: HPAEC-PAD of enzymatic assay after 16 hours on (A) 4Np- α -GlcNAc; (B) 4Np- β -GlcNAc; (C) *N,N*-diacetylchitobiose

The steady state kinetic constants on GlcNAc measured in standard conditions were: K_M 22.01 ± 4.91 mM, k_{cat} 1.94 ± 0.19 s⁻¹ and k_{cat}/K_M $88 \cdot 10^{-3}$ (s⁻¹/mM⁻¹). The low specificity constant suggested that GlcNAc may not be the substrate of SSO2901 *in vivo*. Indeed, for example, *N,N*-diacetyl-chitobiose de-*N*-acetylase Tk-Dac from *T. kodakarensis* displayed a k_{cat}/K_M on *N,N*-diacetyl-chitobiose about 10-fold higher than rSSO2901 on GlcNAc (Tanaka T *et al.*, 2004).

Enzymatic Assays of rSSO2900

As described above, SSO2900 is annotated as a putative UDP-GlcNAc tranferase. To test this hypothesis we incubated 30 µg of rSSO2900 for 16 hours in 50 mM buffer phosphate pH 6.5 at 65 °C with 10 mM UDP-GlcNAc and we run the reaction mixtures on HPAEC-PAD. The chromatograph showed that this enzyme tranferred GlcNAc from UDP-GlcNAc to water, but we could not observe peaks suggesting transglycosylation reactions (Figure 3.4.15). This may suggest that rSSO2900 recognized UDP-GlcNAc as donor, but, in the absence of a suitable acceptor promoted substrate hydrolysis at low efficiency as observed for other GTs (Gómez H *et al.*, 2013). Therefore, both SSO2900 and SSO2901 recognized substrates containing GlcNAc, indicating a probable correlation of these enzymes *in vivo*.

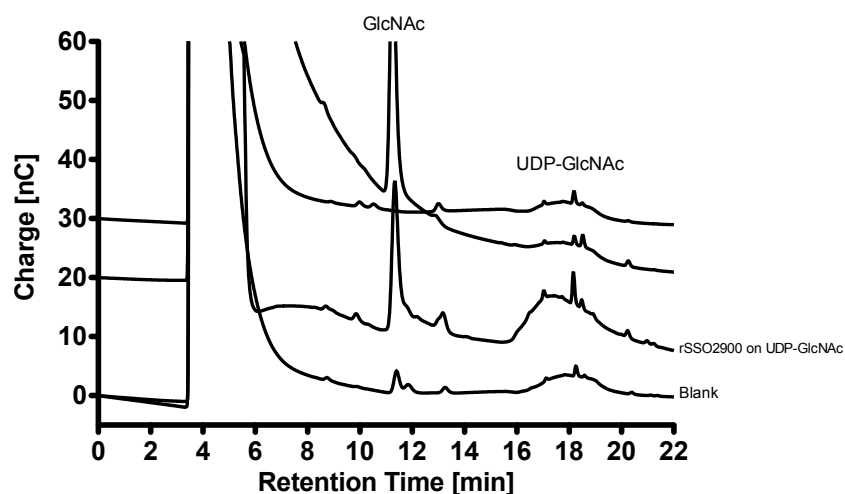


Figure 3.4.15: HPAEC-PAD analysis of rSSO2900 reaction.

In Sulfolobales, β -*N*-acetyl-glucosamine have been identified among the monosaccharides constituents of *N*-glycans of proteins and exopolysaccharides while α -*N*-acetyl-glucosamine is a biosynthesis intermediate of Glycosylphosphatidylinositol (GPI)-anchor (Kobayashi T *et al.*, 1996). The GPI anchor is a post-translational modification composed by complex glycolipids covalently linked to the C-terminus of proteins (Zurzolo C and Simons K, 2015) that provides a stable anchoring to membrane for the proteins (Low MG, 1989; Low MG and Saltiel AR, 1988). Biosynthetic pathway began with the transfer of *N*-acetylglucosamine from UDP-*N*-acetylglucosamine (UDP-GlcNAc) donor to PI by glycosyltransferase. The resulting GlcNAc-PI is deacetylated to yield GlcN-PI (Urbaniak MD and Ferguson MAJ, 2009). The glycosyltransferases that catalyzed the first biosynthesis step, named PIG-A, were classified into family GT4 of CAZy database (Oswal N *et al.*, 2008, Lombard V *et al.*, 2014), while the eukaryotic de-*N*-acetylase homologs, called PIG-L, were not classified in this database. The presence of GPI anchor has been proven in Sulfolobales by Kobayashi and coworkers, but the enzymes involved are still unknown (Kobayashi T *et al.*, 1997). Thus, SSO2901 and SSO2901 might be involved in this biosynthesis pathway. Comparing with the *S. acidocaldarius* genome, SSO2900 and SSO2901

were similar to SUSAZ_08645 (GT4) and SUSAZ_05460 (CE14); sharing 28% and 26% of identity, respectively. However, both *S. acidocaldarius* genes were not characterized, therefore this pathway is still completely unknown in Sulfolobales. In order to shed light on their functions *in vivo*, a more detailed functional and structural characterization is in progress.

In conclusion, our study, which led to the identification of the first unclassified GlcNAc de-*N*-acetylase and a novel UDP-GlcNAc glycosyl transferase from *S. solfataricus* P2, demonstrated that the functional screening of enzymatic activity and the detailed biochemical characterization of the identified enzymes from natural sources is of utmost importance to increase our knowledge on enzyme classes, and to complete the information provided by bioinformatic data banks.

4.5- References

- Albers SV, Meyer BH. (2011) The archaeal cell envelope, *Nat. Rev. Microbiol.* 9 2011 414–426.
- Burda P and Aebi M. (1999) The dolichol pathway of N-linked glycosylation. *Biochim. Biophys. Acta* 1426:239 –257.
- Brock TD, Brock KM, Belly RT, Weiss RL. (1972) *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Mikrobiol.* 1972;84(1):54-68.
- Broadley SG, Gumbart JC, Weber BW, Marakalala MJ, Steenkamp DJ, Sewell BT. (2012) A new crystal form of MshB from *Mycobacterium tuberculosis* with glycerol and acetate in the active site suggests the catalytic mechanism. *Acta Crystallogr D Biol Crystallogr.* 2012 Nov;68(Pt 11):1450-9.
- Calo D, Eilam Y, Lichtenstein RG, Eichler J. (2010) Towards glycoengineering in archaea: replacement of *Haloferax volcanii* AglD with homologous glycosyltransferases from other halophilic archaea. *Appl. Environ. Microbiol.* 2010, 76 (17), 5684–92.
- Cannio R, Di Prizito N, Rossi M, Morana A. (2004) A xylan-degrading strain of *Sulfolobus solfataricus*: isolation and characterization of the xylanase activity. *Extremophiles* 8: 117– 124
- Cardona S, Remonsellez F, Guilianini N, Jerez CA. (2001) The glycogen-bound polyphosphate kinase from *Sulfolobus acidocaldarius* is actually a glycogen synthase. *Applied and environmental microbiology* 67:4773–80.
- Cobucci-Ponzano B, Aurilia V, Riccio G, Henrissat B, Coutinho PM, Strazzulli A, Padula A, Corsaro MM, Pieretti G, Pocsfalvi G, Fiume I, Cannio R, Rossi M, Moracci M (2010b). A new archaeal beta-glycosidase from *Sulfolobus solfataricus*: seeding a novel retaining beta-glycan-specific glycoside hydrolase family along with the human non-lysosomal glucosylceramidase GBA2. *J Biol Chem.* 2010 Jul 2;285(27):20691-703.
- Cobucci-Ponzano B, Conte F, Rossi M, Moracci M. (2008) The alpha-L-fucosidase from *Sulfolobus solfataricus*. *Extremophiles.* 2008 Jan;12(1):61-8.

- Cobucci-Ponzano B, Conte F, Strazzulli A, Capasso C, Fiume I, Pocsfalvi G, Rossi M, Moracci M (2010a). The molecular characterization of a novel GH38 α -mannosidase from the crenarchaeon *Sulfolobus solfataricus* revealed its ability in de-mannosylating glycoproteins. *Biochimie*. 2010 Dec;92(12):1895-907.
- Chang T, Milne KG, Güther ML, Smith TK, Ferguson MA. (2002) Cloning of *Trypanosoma brucei* and *Leishmania major* genes encoding the GlcNAc-phosphatidylinositol de-N-acetylase of glycosylphosphatidylinositol biosynthesis that is essential to the African sleeping sickness parasite. *J Biol Chem*. 2002 Dec 20;277(51):50176-82.
- DeAngelis PL, Gunay NS, Toida T, Mao WJ, Linhardt RJ. (2002) Identification of the capsular polysaccharides of Type D and F *Pasteurella multocida* as unmodified heparin and chondroitin, respectively. *Carbohydr Res*. 2002 Sep 27;337(17):1547-52.
- DeAngelis PL, Liu J, Linhardt RJ. (2013) Chemoenzymatic synthesis of glycosaminoglycans: re-creating, re-modeling and re-designing nature's longest or most complex carbohydrate chains. *Glycobiology*. 2013 Jul;23(7):764-77.
- Deli A, Koutsioulis D, Fadouloglou VE, Spiliotopoulou P, Balomenou S, Arnaouteli S, Tzanodaskalaki M, Mavromatis K, Kokkinidis M, Bouriotis V. (2010) LmbE proteins from *Bacillus cereus* are de-N-acetylases with broad substrate specificity and are highly similar to proteins in *Bacillus anthracis*. *FEBS J*. 2010 Jul;277(13):2740-53.
- Ernst HA, Lo Leggio L, Willemoës M, Leonard G, Blum P, Larsen S. (2006) Structure of the *Sulfolobus solfataricus* alpha-glucosidase: implications for domain conservation and substrate recognition in GH31. *Journal of molecular biology* 358:1106–24.
- Ferrara MC, Cobucci-Ponzano B, Carpentieri A, Henrissat B, Rossi M, Amoresano A, Moracci M. (2014) The identification and molecular characterization of the first archaeal bifunctional exo- β -glucosidase/N-acetyl- β -glucosaminidase demonstrate that family GH116 is made of three functionally distinct subfamilies. *Biochim Biophys Acta*. 2014 Jan;1840(1):367-77.
- Gómez H, Lluch JM, Masgrau L. (2013) Substrate-assisted and nucleophilically assisted catalysis in bovine α 1,3-galactosyltransferase.

- Mechanistic implications for retaining glycosyltransferases. *J Am Chem Soc.* 2013 May 8;135(18):7053-63.
- Grogan DW. (1989) Phenotypic characterization of the archaeobacterial genus *Sulfolobus*: comparison of five wild-type strains. *J. Bacteriol.* 171: 6710 – 6719.
 - Guan Z, Naparstek S, Calo D, Eichler J. (2012) Protein glycosylation as an adaptive response in Archaea: growth at different salt concentrations leads to alterations in *Haloferax volcanii* S-layer glycoprotein N-glycosylation. *Environ. Microbiol.* 2012, 14 (3), 743– 53.
 - Hernick M. (2013) Mycothiol, a target for potentiation of rifampin and other antibiotics against *M. tuberculosis*. *Expert Rev. Anti Infect. Ther.* 2013, 11, 49–67.
 - Hjort K, Bernander R. (2001) Cell cycle regulation in the hyperthermophilic crenarchaeon *Sulfolobus acidocaldarius*. *Mol. Microbiol.* 2001, 40 (1), 225–34.
 - Huang X, Hernick M. (2011) A fluorescence-based assay for measuring N-acetyl-1-D-myo-inositol-2-amino-2-deoxy- α -D-glucopyranoside deacetylase activity. *Anal Biochem.* 2011 Jul 15;414(2):278-81.
 - Jarrell KF, Jones GM, Nair DB. (2010) Biosynthesis and role of N- linked glycosylation in cell surface structures of archaea with a focus on flagella and s layers. *Int. J. Microbiol.* 2010, 470138, 5.
 - Johnson W, Fisher JF.(2013) Mobashery S Bacterial cell-wall recycling. *Ann. N. Y. Acad. Sci.* 1277 2013 54–75.
 - Khor E. (2001) Chitin: Fulfilling a Biomaterials Promise. Ed. Elsevier.
 - Kim MS, Park JT, Kim YW, Lee HS, Nyawira R, Shin HS, Park CS, Yoo SH, Kim YR, Moon TW, Park KH. (2004) Properties of a novel thermostable glucoamylase from the hyperthermophilic archaeon *Sulfolobus solfataricus* in relation to starch processing. *Applied and environmental microbiology* 70:3933–40
 - Kobayashi T, Nishizaki R, Ikezawa H. (1997) The presence of GPI-linked protein(s) in an archaeobacterium, *Sulfolobus acidocaldarius*, closely

related to eukaryotes. *Biochim Biophys Acta*. 1997 Feb 11;1334(1):1-4

- Koerdt A, Gödeke J, Berger J, Thormann KM, Albers SV. (2010) Crenarchaeal biofilm formation under extreme conditions. *PLoS One*. 2010 Nov 24;5(11):e14104.
- Kort JC, Esser D, Pham TK, Noirel J, Wright PC, Siebers B. (2013) A cool tool for hot and sour Archaea: proteomics of *Sulfolobus solfataricus*. *Proteomics*. 2013 Oct;13(18-19):2831-50.
- Lalithambika S, Peterson L, Dana K, Blum P. (2012) Carbohydrate hydrolysis and transport in the extreme thermoacidophile *Sulfolobus solfataricus*. *Appl Environ Microbiol*. 2012 Nov;78(22):7931-8.
- Limauro D, Cannio R, Fiorentino G, Rossi M, Bartolucci S. (2001.) Identification and molecular characterization of an endoglucanase gene, *celS*, from the extremely thermophilic archaeon *Sulfolobus solfataricus*. *Extremophiles* 5: 213– 219.
- Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B, (2014) The carbohydrate-active enzymes database (CAZy) in 2013, *Nucleic Acids Res*. 42 2014 D490–5
- Low MG. (1989) The glycosyl-phosphatidylinositol anchor of membrane proteins. *Biochim Biophys Acta*. 1989 Dec 6;988(3):427-54.
- Low MG, Saltiel AR. (1988) Structural and functional roles of glycosyl-phosphatidylinositol in membranes. *Science*. 1988 Jan 15;239(4837):268-75.
- Maruta K, Mitsuzumi H, Nakada T, Kubota M, Chaen H, Fukuda S, Sugimoto T, Kurimoto M. (1996) Cloning and sequencing of a cluster of genes encoding novel enzymes of trehalose biosynthesis from thermophilic archaeobacterium *Sulfolobus acidocaldarius*. *Biochimica et biophysica acta* 1291:177–81.
- Maurelli L, Giovane A, Esposito A, Moracci M, Fiume I, Rossi M, Morana A. (2008) Evidence that the xylanase activity from *Sulfolobus solfataricus* O α is encoded by the endoglucanase precursor gene (*sso1354*) and characterization of the associated cellulase activity. *Extremophiles* 12: 689– 700

- Meyer BH and Albers SV. (2013) Hot and sweet: protein glycosylation in Crenarchaeota. *Biochem Soc Trans.* 2013 Feb 1;41(1):384-92.
- Moracci M, Cobucci Ponzano B, Trincone A, Fusco S, De Rosa M, van Der Oost J, Sensen CW, Charlebois RL, Rossi M. (2000) Identification and molecular characterization of the first alpha -xylosidase from an archaeon. *J Biol Chem.* 2000 Jul 21;275(29):22082-9.
- Mine S, Ikegami T, Kawasaki K, Nakamura T, Uegaki K. (2012) Expression, refolding, and purification of active diacetylchitobiose deacetylase from *Pyrococcus horikoshii*. *Protein Expression and Purification* 84 2012 265–269.
- Nakamura N, Inoue N, Watanabe R, Takahashi M, Takeda J, Stevens VL, Kinoshita T. (1997) Expression cloning of PIG-L, a candidate N-acetylglucosaminyl-phosphatidylinositol deacetylase. *J Biol Chem.* 1997 Jun 20;272(25):15834-40.
- Newton GL, Av-Gay Y, Fahey RC. (2000) N-Acetyl-1-D-myo-inosityl-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase (MshB) is a key enzyme in mycothiol biosynthesis. *J Bacteriol.* 2000 Dec;182(24):6958-63.
- Newton GL, Buchmeier N, Fahey RC. (2008) Biosynthesis and functions of mycothiol, the unique protective thiol of actinobacteria. *Microbiol. Mol. Biol. Rev.* 2008, 72, 471–494.
- Nicolaus B, Kambourova M, Toksoy Onerc E. (2010) Exopolysaccharides from extremophiles: from fundamentals to biotechnology. *Environmental Technology* Vol. 31, No. 10, September 2010, 1145–1158
- Nicolaus B, Manca MC, Romano I, Lama L (2003) Production of an exopolysaccharide from two thermophilic archaea belonging to the genus *Sulfolobus*. *FEMS Microbiol Lett* 109: 203–206.
- Oswal N, Sahni NS, Bhattacharya A, Komath SS, Muthuswami R. (2008) Unique motifs identify PIG-A proteins from glycosyltransferases of the GT4 family. *BMC Evol Biol.* 2008 Jun 4;8:168. doi: 10.1186/1471-2148-8-168.
- Palmieri G, Balestrieri M, Peter-Katalinic J, Pohlentz G, Rossi M, Fiume I, Pocsfalvi G. (2013) Surface-exposed glycoproteins of hyperthermophilic

Sulfolobus solfataricus P2 show a common N-glycosylation profile, J. Proteome Res. 12 2013 2779–2790.

- Peyfoon E, Meyer B, Hitchen PG, Panico M, Morris HR, Haslam SM, Albers SV, Dell A. (2010) The S-layer glycoprotein of the Crenarchaeote *Sulfolobus acidocaldarius* is glycosylated at multiple sites with chitobiose-linked N-glycans, *Archaea*. 2010 Sep 29;2010.
- Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A, Finn RD. (2012) The Pfam protein families database. *Nucleic Acids Res*. 2012, 40, D290–D301.
- Rolfsmeier M, Blum P. (1995) Purification and characterization of a maltase from the extremely thermophilic crenarchaeote *Sulfolobus solfataricus*. *Journal of bacteriology* 177:482–5.
- Rolfsmeier M, Haseltine C, Bini E, Clark A, Blum P. (1998) Molecular characterization of the alpha-glucosidase gene (*malA*) from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J Bacteriol*. 1998 Mar;180(5):1287-95.
- Santos H, da Costa MS. (2001) Organic solutes from thermophiles and hyperthermophiles. *Methods in Enzymology* 334:302–315.
- Selig M, Xavier KB, Santos H, Schonheit P. (1997) Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways in hyperthermophilic archaea and the bacterium *Thermotoga*. *Arch. Microbiol*. 167:217–232.
- She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC, Clausen IG, Curtis BA, De Moors A, Erauso G, Fletcher C, Gordon PM, Heikamp-de Jong I, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW, Van der Oost J. (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc Natl Acad Sci U S A*. 2001 Jul 3;98(14):7835-40. Epub 2001 Jun 26.
- Smith TK, Crossman A, Borissow CN, Paterson MJ, Dix A, Brimacombe JS, Ferguson MA. (2001) Specificity of GlcNAc-PI de-N-acetylase of GPI biosynthesis and synthesis of parasite-specific suicide substrate inhibitors.

- EMBO J. 2001 Jul 2;20(13):3322-32.
- Tanaka T, Fukui T, Fujiwara S, Atomi H, Imanaka T. (2004) Concerted action of diacetylchitobiose deacetylase and exo-beta-D-glucosaminidase in a novel chitinolytic pathway in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. J Biol Chem. 2004 Jul 16;279(29):30021-7. Epub 2004 May 10.
 - Tripepi M, You J, Temel S, Önder Ö, Brisson D, Pohlschröder M. (2012) N-glycosylation of *Haloferax volcanii* flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. J. Bacteriol. 2012, 194 (18), 4876–87.
 - Urbaniak MD, Ferguson MAJ. (2009) The GlcNAc-PI de-N-acetylase: Structure, function, and activity. Enzymes 2009, 26, 49–64.
 - Viars S, Valentine J, Hernick M. (2014) Structure and function of the LmbE-like superfamily. Biomolecules. 2014 May 16;4(2):527-45.
 - Watanabe R, Ohishi K, Maeda Y, Nakamura N, Kinoshita T. (1999) Mammalian PIG-L and its yeast homologue Gpi12p are N-acetylglucosaminylphosphatidylinositol de-N-acetylases essential in glycosylphosphatidylinositol biosynthesis. Biochem J. 1999 Apr 1;339 (Pt 1):185-92.
 - Whitfield GB, Marmont LS, Howell PL. (2015) Enzymatic modifications of exopolysaccharides enhance bacterial persistence. Front Microbiol. 2015; 6: 471.
 - Worthington P, Hoang V, Perez-pomares F, Blum P. (2003) Targeted Disruption of the alpha-Amylase Gene in the Hyperthermophilic Archaeon *Sulfolobus solfataricus* J Bacteriol. 2003 Jan;185(2):482-8.
 - Yang LL, Haug A. (1979) Purification and partial characterization of a procaryotic glycoprotein from the plasma membrane of *Thermoplasma acidophilum*. Biochim. Biophys. Acta 556:265–277.
 - Yurist-Doutsch S, Magidovich H, Ventura VV, Hitchen PG, Dell A, Eichler J. (2010) N-Glycosylation in Archaea: on the coordinated actions of *Haloferax volcanii* AglF and AglM. Mol. Microbiol. 2010, 75 (4), 1047–58.
 - Zillig W, Stetter KO, Wunderl S. (1980) The *Sulfolobus*-“*Caldariella*” group: taxonomy on the basis of the structure of DNA-dependent RNA

polymerases. Arch. Microbiol. 125:259 –269.

- Zolghadr B, Klingl A, Koerdt A, Driessen AJM, Rachel R, Albers SV. (2010) Appendage-mediated surface adherence of *Sulfolobus solfataricus*. J Bacteriol 192:104–110
- Zou Y, Brunzelle JS, Nair SK. (2008) Crystal structures of lipoglycopeptide antibiotic deacetylases: implications for the biosynthesis of A40926 and teicoplanin. Chem Biol. 2008 Jun;15(6):533-45.
- Zurzolo C, Simons K. (2016) Glycosylphosphatidylinositol-anchored proteins: Membrane organization and transport. Biochim Biophys Acta. 2016 Apr;1858(4):632-9.

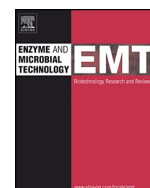
~~~~~Publication~~~~~

Cobucci-Ponzano B, Strazzulli A, **Iacono R**, Masturzo G, Giglio R, Rossi M, Moracci M. *Novel thermophilic hemicellulases for the conversion of lignocellulose for second generation biorefineries*. Enzyme Microb Technol. 2015 Oct;78:63-73. doi: 10.1016/j.enzmictec.2015.06.014. Epub 2015 Jun 25.



Contents lists available at ScienceDirect

Enzyme and Microbial Technology

journal homepage: www.elsevier.com/locate/emt

Novel thermophilic hemicellulases for the conversion of lignocellulose for second generation biorefineries

Beatrice Cobucci-Ponzano, Andrea Strazzulli, Roberta Iacono, Giuseppe Masturzo, Rosa Giglio, Mosè Rossi, Marco Moracci*

Institute of Biosciences and Bioresources, National Research Council of Italy, Via P. Castellino 111, 80131 Naples, Italy

ARTICLE INFO

Article history:

Received 23 April 2015

Received in revised form 13 June 2015

Accepted 20 June 2015

Available online 25 June 2015

Keywords:

Bioconversion

Biocatalysis

Biotransformation

Biofuel

White chemistry

Thermophilic enzymes

ABSTRACT

The biotransformation of lignocellulose biomasses into fermentable sugars is a very complex procedure including, as one of the most critical steps, the (hemi) cellulose hydrolysis by specific enzymatic cocktails. We explored here, the potential of stable glycoside hydrolases from thermophilic organisms, so far not used in commercial enzymatic preparations, for the conversion of glucuronoxylan, the major hemicellulose of several energy crops. Searches in the genomes of thermophilic bacteria led to the identification, efficient production, and detailed characterization of novel xylanase and α -glucuronidase from *Alicyclobacillus acidocaldarius* (GH10-XA and GH67-GA, respectively) and a α -glucuronidase from *Caldicellulosiruptor saccharolyticus* (GH67-GC). Remarkably, GH10-XA, if compared to other thermophilic xylanases from this family, coupled good specificity on beechwood xylan and the best stability at 65 °C (3.5 days). In addition, GH67-GC was the most stable α -glucuronidases from this family and the first able to hydrolyse both aldouronic acid and aryl- α -glucuronic acid substrates. These enzymes, led to the very efficient hydrolysis of beechwood xylan by using 7- to 9-fold less protein (concentrations <0.3 μ M) and in much less reaction time (2 h vs 12 h) if compared to other known biotransformations catalyzed by thermophilic enzymes. In addition, remarkably, together with a thermophilic β -xylosidase, they catalyzed the production of xylose from the smart cooking pre-treated biomass of one of the most promising energy crops for second generation biorefineries. We demonstrated that search by the CAZy Data Bank of currently available genomes and detailed enzymatic characterization of recombinant enzymes allow the identification of glycoside hydrolases with novel and interesting properties and applications.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Glycosidases extracted from thermophilic organisms show many advantages over biocatalysts working at temperatures <50 °C and demonstrated remarkable utility in the bioconversion of car-

bohydrates by enzymes highly resistant to temperature, chemicals, and pH [1–5]. In particular, sustainable chemo-enzymatic lignocellulose conversions for the production of biofuels and precursors of fine chemicals [6–10] can take great advantage by hyperstable glycoside hydrolases [9,11–15].

In fact, a still critical aspect of lignocellulosic-derived second generation plants, exploiting the so-called energy crops as an alternative to crops used for human and animal food production, is the composition of the enzymatic cocktails specifically tailored for the biomass of choice (for reviews see [7,8,10,16,17]). A specific example is the degradation of glucurono-xylan [18], the major hemicellulose of *Arundo donax* (giant reed) a fast growing and low input-high yielding energy crop [19].

The enzymatic degradation of this linear polymer of β -1,4-linked D-xylopyranosyl units frequently substituted with α -1,2-linked glucuronate residues, which are often methylated on the C4 (MeGlcA) [20] occurs by the concerted action of xylanases and α -glucuronidases; thus, we embarked in the search for novel enzymes of this kind in the genomes of (hyper) thermophilic microorganisms

Abbreviations: CAZy, carbohydrate active enzyme; 3D, three-dimensional; MGX, 4-O-methyl-glucurono-xylan; 2Np-Cel, 2-nitrophenyl- β -cellobioside; 4Np-Xyl, 4-nitrophenyl- β -xylopyranoside; 4Np-Glc, 4-nitrophenyl- β -glucopyranoside; CMC, carboxy-methyl-cellulose; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TLC, thin layer chromatography; ORF, open reading frame; 4Np-GluA, 4-nitrophenyl- α -glucuronide; XOs, xylooligosaccharides; Xyl1, xylose; Xyl2, xylobiose; Xyl3, xylotriose; Xyl4, xylo-tetraose; Xyl5, xylopentaose; Xyl6, xylohexaose; HPAEC-PAD, high performance anionic exchange chromatography with pulsed amperometric detection; MeGlcA, 4-O-methyl glucuronic acid.

* Corresponding author. Fax: +39 0816132634.

E-mail addresses: beatrice.cobucciponzano@ibbr.cnr.it (B. Cobucci-Ponzano), andrea.strazzulli@ibbr.cnr.it (A. Strazzulli), roberta.iacono@ibbr.cnr.it (R. Iacono), giuseppe.masturzo@hotmail.com (G. Masturzo), rosa.giglio@ibbr.cnr.it (R. Giglio), mose.rossi@ibbr.cnr.it (M. Rossi), marco.moracci@ibbr.cnr.it (M. Moracci).

<http://dx.doi.org/10.1016/j.enzmictec.2015.06.014>

0141-0229/© 2015 Elsevier Inc. All rights reserved.

which are known for containing unique (hemi) cellulolytic systems [21–26].

α -Glucuronidases (EC 3.2.1.131 and .139), belonging to families GH4, GH67 and GH115 of carbohydrate active enzyme (CAZy) classification (www.cazy.org), are still not extensively studied with about 2, 20 and 4 members characterized from these families, respectively, [27]. Family GH4 contains several enzymatic activities and do not hydrolyze 4-O-methyl- α -glucuronoxylan or its oligosaccharidic fragments [28]. Instead, GH67 and GH115 enzymes, including only α -glucuronidases, show the formers selectively removing the MeGlcA-1,2 bound to the non-reducing end xylose of short oligosaccharides of glucurono-xylans [29,30]. Instead, GH115 enzymes remove glucuronic acid from both the terminal and the internal regions of xylooligosaccharides and xylans [31–33]. A more detailed inspection of these families revealed that enzymes from thermophilic microorganisms are concentrated in GH67 with only the uncharacterized enzyme from *Thermobispora bispora* belonging to GH115. Therefore, we focused our search among the thermophilic GH67 enzymes acting on MeGlcA bound to the non-reducing end of xylooligosaccharides.

Xylanases (EC 3.2.1.8 and .32) are currently classified in families GH5, GH8, GH10, GH11, GH26, GH30, GH43 and GH51. Among these, GH10 and GH11 enzymes, which are endo- β -xylanases whose catalytic apparatus and three-dimensional structures are well known, cleave glucuronoxylan chains when MeGlcA is linked to xylose at the +1 and +2 subsites, respectively [27,34,35]. Therefore, GH10 products, showing MeGlcA at the non-reducing end, are direct substrates of α -glucuronidases, making the xylanases from this family the ideal candidates for hydrolysis of glucuronoxylan in cooperation with GH67 enzymes.

Here, we show the exploitation of glycoside hydrolases belonging to families GH10 and GH67 from the thermophilic bacteria *Alicyclobacillus acidocaldarius* and *Caldicellulosiruptor saccharolyticus*, which contain a large survey of carbohydrate active enzymes proved to be useful for glycan biotransformations [36–40]. We demonstrate that these enzymes, which showed remarkable stability to high temperatures and unique substrate specificities, efficiently hydrolyzed (methyl)-glucurono xylans and smart cooking pre-treated *A. donax* biomass.

2. Materials and methods

2.1. Materials

All commercially available substrates were purchased from Sigma–Aldrich, Carbosynth and Megazyme. The synthetic oligonucleotides were from PRIMM (Italy). The biomass of *A. donax* used in this study derived from a pre-treatment step (smart cooking) described in patent WO2010113129A2. This treatment comprised the soaking of the ligno-cellulosic biomass feedstock in vapor or liquid water or mixture thereof in the temperature range of 100–210 °C for 1 min to 24 h to create a soaked biomass containing a dry content and a first liquid. Steam exploding of the former created a steam exploded stream comprising solids and a second liquid, which was used for the experiments described. This treatment was able to break up the lignocellulosic matrix and avoid the formation of inhibitor compounds using only steam and water without the requirement of additional chemical products.

2.2. Cloning and purification

The Aaci.2328 gene, coding GH10-XA xylanase, was amplified by PCR from the genome of *A. acidocaldarius* ATCC27009 using the synthetic oligonucleotides 2328-5' (5'-CACCATGACGGATCAAGCGCGCT-3') and 2328-3' (5'-

GTTTTGGGCGAGGCGCACCAC-3'). The amplification reaction was performed with the PfuUltra II Fusion HS DNA Polymerase (Stratagene) by using the following program: hot start 5 min at 95 °C; 5 cycles 1 min at 95 °C, 1 min at 50 °C and 1.5 min at 72 °C; 30 cycles 1 min at 95 °C, 1 min at 60 °C, and 1.5 min at 72 °C; final extension 10 min at 72 °C. The DNA fragment obtained was cloned in the expression vector pET101/D-TOPO (Invitrogen), obtaining the recombinant plasmid pET101/D-TOPO-Aaci.2328. Here, the gene is under the control of an isopropyl-1-thio- β -D-galactopyranoside (IPTG) inducible T7 RNA polymerase promoter and the C-terminal of the protein was fused to V5 epitope and 6xHis tag. PCR-generated construct was verified by sequencing and the ORF was expressed in *Escherichia coli* cells, strain BL21 Star (DE3) (Invitrogen), according to the manufacturer. The cells transformed with pET101/D-TOPO-Aaci.2328 were grown at 37 °C in 2 L of Luria–Bertani (LB) broth supplemented with ampicillin (50 μ g mL⁻¹). Gene expression was induced by the addition of 0.1 mM IPTG when the culture reached an A₆₀₀ of 1.0. Growth was allowed to proceed for 16 h, and cells were harvested by centrifugation at 5000 \times g. The resulting cell pellet was resuspended in 50 mM sodium phosphate buffer, pH 8.0, 300 mM NaCl and 1% TRITON-X100 with a ratio of 5 mL g⁻¹ cells and then was incubated at 37 °C for 1 h with 20 mg of lysozyme (Fluka) and 25 U g⁻¹ cell of Benzonase (Novagen). Cells were lysed by French cell pressure treatment and cell debris were removed by centrifugation at 12,000 \times g for 30 min. The free cellular extract (FCE) was loaded on a His Trap FF crude column (GE-Healthcare) equilibrated with 50 mM sodium phosphate buffer, pH 8.0, 300 mM NaCl (Buffer A). After an initial wash-step (20-column volumes) with buffer A, the protein was eluted with a two-step gradient of imidazole in Buffer A (250 mM imidazole, 20-column volumes followed by 500 mM imidazole, 20-column volumes) at a flow rate of 1 mL min⁻¹. The protein was eluted at 250 mM imidazole. The active fractions were pooled, dialyzed against 20 mM sodium phosphate buffer, pH 7.3, 150 mM NaCl (PBS buffer) and then heat-fractionated for 30 min at 50 °C and, after centrifugation, for additional 30 min at 60 °C. The resulting supernatant was stored at 4 °C. The protein concentration was determined with the Bradford assay [41]. After this procedure GH10-XA was more than 95% pure by SDS-PAGE.

The Aaci.0060 gene, coding GH67-GA α -glucuronidase, was amplified by PCR from the genome of *A. acidocaldarius* ATCC27009 using the synthetic oligonucleotides 0060-5' (5'-CACCCTTGACGAACATCCCTGA-3') and 0060-3' (5'-CGGGTAGATGTGGAGCCC-3'). The amplification reaction was performed with the PfuUltra II Fusion HS DNA Polymerase (Stratagene) by using the following program: hot start 5 min at 95 °C; 5 cycles 1 min at 95 °C, 1 min at 48 °C and 3 min at 72 °C; 30 cycles 1 min at 95 °C, 1 min at 55 °C, and 3 min at 72 °C; final extension 10 min at 72 °C. The DNA fragment obtained was cloned in the expression vector pET101/D-TOPO (Invitrogen), as described above, obtaining the recombinant plasmid pET101/D-TOPO-Aaci.0060. The Csac.2689 gene, coding GH67-GC α -glucuronidase, was amplified by PCR from the genome of *C. saccharolyticus* DSM8903 using the synthetic oligonucleotides 2689-5' (5'-CACCATGGAACACGTCAAACAAAAA-3') and 2689-3' (5'-TGGATATATAAGTCTTCTTTTTCATC-3'). The amplification reaction was performed with the Platinum Taq High Fidelity (Invitrogen) by using the following program: hot start 5 min at 94 °C; 5 cycles 1.5 min at 94 °C, 1 min at 50 °C and 3 min at 68 °C; 30 cycles 1.5 min at 94 °C, 1 min at 55 °C, and 3 min at 68 °C; final extension 10 min at 68 °C. The DNA fragment obtained was cloned in the expression vector pET101/D-TOPO (Invitrogen), as described above, obtaining the recombinant plasmid pET101/D-TOPO-Csac.2689. The ORFs were expressed in *E. coli* cells as described above, and gene expression was induced by the addition of 0.5 mM IPTG. After harvesting, the resulting cell pellets were resuspended,

lysed and loaded on a His Trap FF crude column (GE-Healthcare) as described above. Both proteins were eluted at 250 mM imidazole. The active fractions were pooled, dialyzed against PBS buffer, and then heat-fractionated for 30 min at 50 °C for GH67-GA and for 30 min at 50 and 60 °C and 20 min at 70 °C for GH67-GC. The resulting supernatants of GH67-GA and GH67-GC were loaded on a HiLoad 16/60 Superdex column (Amersham Bioscience) equilibrated in PBS buffer. The active fractions were pooled and stored at 4 °C. The protein concentration was determined with the Bradford assay [41]. After this procedure, GH67-GA and GH67-GC were 80 and 95% pure by SDS-PAGE, respectively.

The β -xylosidase from *Thermotoga thermarum* (GH3-XT) was cloned, expressed and purified as described elsewhere [42] with some modifications. Plasmid pET-20b-GH3-XT was transformed into *E. coli* BL21 (DE3), and induced to express recombinant β -xylosidase by adding IPTG to final concentration of 0.5 mM at OD600 approximately 0.8, and incubated further at 37 °C for about 16 h. 2 L of the recombinant cells carrying pET-20b-GH3-XT were harvested by centrifugation (6000 \times g, 15 min, 4 °C), and resuspended in 20 mL of 5 mM imidazole, 0.5 mM NaCl, and 20 mM Tris-HCl buffer (pH 7.9). The cell extracts after sonication were heat treated (70 °C, 30 min), and then cooled in an ice bath, and centrifuged (15,000 \times g, 4 °C, 20 min). The obtained supernatants were loaded onto a His Trap FF Crude (1 mL) (GE-Healthcare) with a flow rate 1 mL min⁻¹. Finally, 1 mL fractions were collected by eluting with 0.4 M imidazole, 0.5 M NaCl, and 20 mM Tris-HCl buffer (pH 7.9). The fractions showing β -xylosidase activity were dialysed against PBS 1X, pH 7.3.

2.3. Molecular mass determination

The molecular mass of the three enzymes was determined by gel filtration on a Superdex 200HR 10/300 FPLC column (GE-Healthcare) performed as described above. Molecular weight markers were apoferritin (443 kDa), amylase (200 kDa), albumin (66 kDa) and ribonuclease A (14 kDa).

2.4. Standard assays

The standard assay for GH10-XA was performed on 5 mg mL⁻¹ of beechwood xylan in 50 mM sodium phosphate buffer at pH 6.5 at 65 °C by using 1 μ g of enzyme in the final volume of 0.1 mL. The standard assay for GH67-GA was performed on 2 mg mL⁻¹ of aldouronic acids mixture (Megazyme) in 100 mM sodium phosphate buffer at pH 6.5 at 50 °C by using 17 μ g of enzyme in the final volume of 0.1 mL. The standard assay for GH67-GC was performed on 1 mg mL⁻¹ of aldouronic acids mixture (Megazyme) in 100 mM sodium phosphate buffer at pH 6.0 at 65 °C by using 10 μ g of enzyme in the final volume of 0.1 mL. The relative activity was measured by the Somogyi–Nelson assay [43,44], estimating the amount of reducing sugars released after 1 min. One unit of activity was defined as the amount of enzyme releasing 1 μ mol of reducing equivalents per minute at the conditions described.

2.5. Temperature and pH influence

The temperature and pH optima were determined by assaying GH10-XA in 50 mM and GH67 enzymes in 100 mM of the indicated buffers at different pHs in the range of 40–85 °C using the standard assay conditions of each enzyme as reported elsewhere [45]. Thermal stability was evaluated by incubating the enzymes in PBS buffer, at the indicated temperatures. At intervals, aliquots were withdrawn, transferred in ice and assayed at standard conditions. The residual activities were expressed as a percentage of the maxi-

mal enzymatic activity measured before the incubation at indicated temperatures.

2.6. Enzymatic characterization

Substrate specificity and mode of action. The activity of GH10-XA (1 μ g) was tested on MGX (5 mg mL⁻¹), CMC (5 mg mL⁻¹), Avicel (5 mg mL⁻¹), cellobiose (8 mM), and cellotriose (8 mM), at standard conditions; the amount of reducing ends released where measured by Somogyi–Nelson assay as described above. Moreover, the activity of GH10-XA (3 μ g) was tested on 1 mg/mL of XO₅ (from Xyl₂ up to Xyl₅) for 16 h at standard conditions. The products were analyzed on silica gel 60 F254 TLC by using *n*-butanol/methanol/water (50:25:25 v/v) as eluent. The activity on 2Np-Cel (10 mM), 4Np-Glc (5 mM) and 4Np-Xyl (8 mM) was tested in standard conditions. Transglycosylation reactions were performed by incubating GH10-XA (3 μ g) at standard conditions with 30 mM 4Np-Xyl and 8 mM 2Np-Cel in a final volume of 0.2 mL. At time intervals (up to 16 h), aliquots of the reaction mixtures were analyzed on a silica gel 60 F254 TLC by using ethyl acetate/methanol/water (70:20:10 v/v) as eluent. The activity of GH67-GA (17 μ g) was tested on aldouronic acids (2 mg mL⁻¹) in standard conditions; GH67-GC (10 μ g) was tested on 4Np-GlcUA (0.7 mM) and Aldouronic acids (2 mg mL⁻¹) in standard conditions. At 405 nm the molar extinction coefficients of 2-nitrophenol and 4-nitrophenol were 2740 M⁻¹ cm⁻¹ and 9340 M⁻¹ cm⁻¹, respectively. In all the assays, identical mixtures containing all the reagents but the enzyme were prepared as control. To analyze the activity of GH67-GC on aldouronic acids, aliquots of the reaction were analyzed by TLC, as described above, by using *n*-butanol/methanol/water (50:25:25 v/v) as eluent. All the TLCs were detected by exposure to 4% α -naphthol in 10% sulphuric acid in ethanol followed by charring. The reactions on aldouronic acids were also analyzed by HPAEC-PAD on a CarboPac PA200 Analytical (2 \times 205 mm) (Dionex, 1996) by using the following program: flow 0.4 mL min⁻¹, 35 °C, isocratic 100 mM NaOH, [segment 1] 0–25' up to 125 mM CH₃COONa, [segment 2] 25–31.5', up to 450 mM CH₃COONa, [segment 3] up to 35', 450 mM CH₃COONa.

Steady-state kinetic constants. Kinetic constants of GH10-XA on beechwood xylan and on MGX were measured at standard conditions by using the substrates ranging from 0.5 to 25 mg mL⁻¹ and from 0.3 to 7 mg mL⁻¹, respectively and determining the amount of reducing sugars with the Somogyi–Nelson assay as described above. Kinetic constants on the chromogenic substrates 2Np-Cel and 4Np-Xyl were measured at standard conditions by using concentrations of substrate ranging between 0.01–25 and 1–35 mM, respectively. Kinetic constants of GH67-GA and GH67-GC on aldouronic acids were measured at standard conditions by using the substrate ranges 0.05–2.5 and 0.02–1 mg mL⁻¹, respectively. Kinetic constants of GH67-GC on 4Np-GluA by using concentrations of substrate ranging between 0.3 and 7 mM. In all the assays, one unit of enzyme activity was defined as the amount of enzyme catalyzing the hydrolysis of 1 μ mol of substrate in 1 min at the conditions described. Spontaneous hydrolysis of the substrates was subtracted by using appropriate blank mixtures without enzyme. All kinetic data were calculated as the average of at least two experiments and were plotted and refined with the program Prism 5.0 (GraphPad Software, San Diego, California, USA).

2.7. Synergic action of GH10-XA, GH67-GA and GH67-GC

To test the effect of simultaneous incubation of the two enzymes, GH10-XA (0.5 U; 0.26 μ M) and GH67-GC (0.02 U; 0.09 μ M) or GH67-GA (0.07 U; 0.11 μ M) were incubated in 100 mM sodium phosphate buffer, pH 6.0, at 50 °C in the presence of MGX (5 mg mL⁻¹) for 16 h. In addition, the simultaneous action of

GH10-XA and GH67-GC (at the same conditions shown above) was analyzed incubating the two enzyme in 100 mM sodium phosphate buffer, pH 6.0, at 65 °C in the presence of MGX (0.5%; 5 mg mL⁻¹) and beechwood xylan (1.5%, 15 mg mL⁻¹) for 2 h. The amount of reducing ends in 100 µL of reaction mixture was measured with the Somogyi–Nelson assay (see above). A blank mixture with no enzymes and three aliquots with GH10-XA, GH67-GC and GH67-GA alone were used as controls. The reaction products of GH10-XA and GH67-GC were analyzed by HPAEC-PAD on a CarboPac PA200 Analytical (2 × 205 mm) as described above.

GH10-XA and GH67-GC (at the same conditions shown above) and GH3-XT (0.37 U; 0.08 µM) were assayed on MGX (5 mg mL⁻¹) and beechwood xylan (15 mg mL⁻¹) for 2 h and the amount of reducing ends were measured as reported above.

The activities of GH10-XA, GH67-GC and GH3-XT were tested on the liquid fraction of smart cooking pre-treated *A. donax*. For the assays, the biomass was diluted 3-fold with water and dialysed (cut-off 1 kDa) against water. GH10-XA (14 U; 5.1 µM), GH67-GC (0.1 U; 0.32 µM) and GH3-XT (7 U; 0.76 µM) were incubated in 50 mM sodium phosphate buffer, pH 6.0, at 65 °C in the presence of 0.25 mL of biomass prepared as described above. The reaction products were analyzed by HPAEC-PAD on a CarboPac PA200 Analytical (2 × 205 mm) by using the following program: flow 0.25 mL min⁻¹, 35 °C, [segment 1] 0–20' isocratic 20 mM NaOH, [segment 2] 20–40' up to 1 M CH₃COONa, [segment 3] up to 50' 1 M CH₃COONa. Fucose (0.5 nmol) was added as internal standard.

3. Results and discussion

Cloning and expression of GH10 and GH67 cazymes. The genome of the thermophilic bacterium *A. acidocaldarius* ATCC27009 contains the 1017 bp long open reading frame (ORF) Aaci.2328 [GenBank at the NCBI ACV59335.1] encoding for a putative xylanase of 338 amino acids, named GH10-XA, that has never been characterized so far. The amino acid sequence had low identity (62%) to a characterized GH10 xylanase (Additional Table 1) [46], indicating that GH10-XA might show different structural characteristics and substrate specificity. The presence of a possible signal peptide at the amino-terminal of this putative xylanase was excluded by using the prediction program PRED-TAT (Additional Fig. 1A). In addition, the program SignalP 4.0 did not reveal any proteolytic signalling sequence, suggesting that this protein was intracellular (Additional Fig. 1B).

Supplementary material related to this article found, in the online version, at <http://dx.doi.org/10.1016/j.enzmictec.2015.06.014>

The genomes of *A. acidocaldarius* ATCC27009 and of another thermophilic bacterium, *C. saccharolyticus* DSM8903, contain the ORFs Aaci.0060 [GenBank at the NCBI ACV57124.1] and Cscac.2689 [GenBank at the NCBI ABP68258.1] encoding two putative GH67 α-glucuronidases of 690 and 696 amino acids, respectively, and 49% identical, which have never been characterized. These enzymes, named GH67-GA and GH67-GC, have, respectively, up to 57% and 77% amino acid sequence identity to GH67 enzymes characterized so far (Additional Table 2) [47,48]. Also for these enzymes, the analyses with the programs PRED-TAT (Additional Fig. 2A and C) and SignalP 4.0 (Additional Fig. 2B and D) led us to exclude signal peptides for secretion suggesting that they, like GH10-XA, are intracellular.

Supplementary material related to this article found, in the online version, at <http://dx.doi.org/10.1016/j.enzmictec.2015.06.014>

The sequence of the genes amplified from *A. acidocaldarius* ATCC27009 revealed some differences if compared to that reported in the database. In particular, the sequence of GH10-XA differs in 13 nucleotides; however, only one mutation determined an aminoacidic substitution (Gly512Asp), which has been confirmed

in two independent clones and is conserved in homolog enzymes from different *A. acidocaldarius* strains. Instead, the sequence of GH67-GA differ from the one deposited in GenBank at the NCBI in about 150 nucleotides, determining 34 aminoacidic substitutions, 12 of them conserved in the *A. acidocaldarius* Tc4-1 strain. The DNA sequences of GH10-XA and GH67-GA determined here were deposited in GenBank at the NCBI with the accession numbers KJ466334 and KJ466335, respectively.

All three genes were cloned in frame with a sequence encoding a 6His-tag at the carboxy-terminal of the expressed proteins, greatly facilitating their purification by affinity chromatography (Additional Table 3). Subsequent heating fractionations and, in the case of the GH67 enzymes, a S-200 gel filtration chromatography, allowed to remove contaminating *E. coli* proteins and to produce the recombinant enzymes with purity grade estimated by SDS-PAGE of 95%, 80%, and 99% for GH10-XA, GH67-GA and GH67-GC respectively (Additional Fig. 3A–C). Good purification yields (30%) were obtained for GH10-XA and GH67-GC (6.0 mg and 3.1 mg L⁻¹ of bacterial culture, respectively), while GH67-GA resulted in only 0.6 mg L⁻¹ of bacterial culture with a lower purification yield (10%) (Additional Table 3).

Supplementary material related to this article found, in the online version, at <http://dx.doi.org/10.1016/j.enzmictec.2015.06.014>

GH10 and GH67 characterization. The molecular mass of the recombinant GH10-XA in denaturing conditions was about 40 kDa (Additional Fig. 3A), confirming the 41,845 Da deduced from the sequence with the additional His-tag. Size exclusion chromatography in native conditions showed a molecular mass of 39 ± 2 kDa indicating a monomeric structure (Additional Fig. 4A). GH67-GA and GH67-GC show an identical molecular mass of about 80 kDa in denaturing conditions (Additional Fig. 3B and C), fitting well the predicted molecular mass deduced from the sequences of 79,963 and 83,609 Da, respectively. The gel filtration in native conditions revealed molecular masses of 153 ± 2 and 109 ± 18 kDa for GH67-GA and GH67-GC, respectively, which might indicate either relaxed monomeric or, by analogy with all the other GH67 enzymes, tight dimeric structures (Additional Fig. 4B and C) [49].

Supplementary material related to this article found, in the online version, at <http://dx.doi.org/10.1016/j.enzmictec.2015.06.014>

The enzymatic activity of GH10-XA was tested on 5 mg mL⁻¹ beechwood xylan (see Methods). The enzyme is optimally active in the pH range 6.0–8.0 and at 75 °C (Fig. 1A and B) and maintains more than 50% activity in 50 mM sodium phosphate buffer, pH 9.0, at 65 °C (Fig. 1A). The activity at pH near the neutrality further indicates that GH10-XA is an intracellular enzyme as *A. acidocaldarius* grows optimally at pH 3.0–4.0. GH10-XA showed a remarkable stability being still fully active after 3.5 days of incubation at 65 °C and maintaining 40% specific activity after 1 h at 75 °C (Fig. 1C). It is worth noting the remarkable activation observed after 10 min at 65 °C, a phenomenon observed also in other thermophilic enzymes [50]. From these data, all the following characterization of GH10-XA was performed in 50 mM sodium phosphate buffer pH 6.5 at 65 °C. The above characterization indicates a remarkable stability of GH10-XA making this enzyme a promising tool for several potential applications in a wider variety of conditions. In fact, the GH10 xylanase with the highest sequence identity to GH10-XA, from a different *A. acidocaldarius* strain, showed maximal activity at 55 °C and was stable only 20 min at 65 °C [46] (Table 1). The specific activity of this enzyme on beechwood xylan was not reported, precluding a direct comparison with GH10-XA that, on this substrate, is more active than the enzymes from *Thermotoga thermarum*, *Volvariella volvacea*, *Thermoanaerobacterium saccharolyticum*, and *C. lactoaceticus*. In addition, GH10-XA is more stable than the enzymes from *Streptomyces olivaceoviridis*, *Aspergillus niger*, and *Bispora*, *Streptomyces* and *Actinomyces* spp, which show higher specific activity on beechwood xylan (Table 1).

Table 1

Comparison of characterized thermophilic GH10 xylanases.

Microorganism	Mw (kDa)	Optimum temperature (°C)	Optimum pH	Specific activity (U/mg)	Xylan substrate	Reference	Thermostability	Identity vs GH10-XA (%)
<i>Alicyclobacillus acidocaldarius</i> ATCC27009	39	75	6.5	152	Beechwood	This study	>100% (65 °C/3.5 days) –40% (75 °C/1 h) –10% (85 °C/5 m)	100
<i>Alicyclobacillus acidocaldarius</i> sp. A4	42.5	55	7.0	420	Birchwood	[46]	90% (60 °C/1 h) –40% (65 °C/1 h) –0% (70 °C/10 m)	62
<i>Caldibacillus cellulovorans</i>	57	90	6.0	256	Birchwood	[51]	85% (70 °C/4 h) –50% (80 °C/35 m)	43
<i>Thermotoga thermarum</i>	131	95	7.0	146	Beechwood	[42]	100% (85 °C/2 h) –30% (90 °C/2 h) –0% (95 °C/60 m)	41
<i>Geobacillus stearothermophilus</i> T-6	43	70	6.5	280	Oat spelt	[52]	100% (65 °C/10 h) –65% (70 °C/10 h) –50% (75 °C/20 m)	40
<i>Thermotoga maritima</i>	35	105	6.2	126.3	Oat spelt	[53]	50% (90 °C/90 m) –50% (100 °C/8 m) –50% (105 °C/2 m)	36
<i>Streptomyces olivaceoviridis</i> E-86	1,200	60	6.0	332	Beechwood	[54]	50% (65 °C/30 m)	33
<i>Volvariella volvacea</i>	39	60	7.0	67	Beechwood	[55]	70% (50 °C/60 m) –10% (55 °C/60 m)	23
<i>Aspergillus niger</i>	33	60	5.0	3,200	Beechwood	[56]	90% (60 °C/30 m) –15% (70 °C/10 m) –5% (80 °C/5 m)	21
<i>Bispora</i> sp MEY-1	70	60	3.0	2,463	Beechwood	[57]	90% (60 °C/1 h) –25% (70 °C/10 m)	18
<i>Thermoanaerobacterium saccharolyticum</i> NT0U1	50	63	6.4	78	Beechwood	[58]	65% (60 °C/2 h) –30% (65 °C/2 h) –10% (75 °C/10 m)	16
<i>Caldicellulosiruptor lactoaceticus</i>	47	80	6.5	44	Beechwood	[59]	100% (75 °C/6 h) –70% (80 °C/6 h) –25% (85 °C/30 m)	16
<i>Syncephalastrum racemosum</i> Cohn	29	50	8.5	1,402	Birchwood	[60]	not reported	Native
<i>Actinomadura</i> sp. strain Cpt20	20	80	10.0	712	Beechwood	[61]	100% (60 °C/10 h) –30% (70 °C/10 h) –0% (80 °C/9 h)	Native
<i>Streptomyces</i> sp. CS428	37	80	7.0	926,103	Beechwood	[62]	90% (50 °C/1 h) –60% (60 °C/1 h) –50% (70 °C/1 h)	Native
<i>Bacillus</i> sp. AR-009	48	70–75	9.0–10.0	367.9	oat spelt	[63]	74% (60 °C/3 h) –67% (65 °C/1 h)	Native

The pH dependence of GH67-GA and GH67-GC was analysed at 50 °C and 65 °C, respectively, by using aldouronic acids as substrate. The optimal pHs of 6.0 and 6.5 were observed for GH67-GA and GH67-GC, respectively, in 100 mM sodium phosphate buffer (Fig. 2A). Both enzymes were thermophilic, but whereas GH67-GC was optimally active at 65 °C and maintained 80% specific activity at 80 °C, the maximal activity of GH67-GA was observed at 50 °C (Fig. 2B). In addition, the latter enzyme resulted rather unstable to heat, losing 50% of its activity after 30 min at 50 °C. In contrast, it is worth mentioning that the remarkable stability of GH67-GC, which maintained 50% activity after 16 h and 2 h at 65 °C and 75 °C, respectively (Fig. 2C), makes it one of the most stable thermophilic α -glucuronidases known (Table 2). On the basis of these results, the following characterization was performed in 100 mM sodium phosphate buffer pH 6.5, 65 °C for GH67-GC and pH 6.0, 60 °C for GH67-GA.

GH10 and GH67 substrate specificity. The substrate specificity of GH10-XA was tested on 4-O-methyl-glucurono-xylan (MGX), beechwood xylan, 2-nitrophenyl- β -cellobioside (2Np-Cel), 4-nitrophenyl- β -xylopyranoside (4Np-Xyl), 4-nitrophenyl-

β -glucopyranoside (4Np-Glc), carboxy-methyl-cellulose (CMC), Avicel, cellobiose, and celotriose at standard conditions (Additional Table 4). The enzyme was specific for xylose containing polysaccharides, but also remarkably active on 2Np-Cel while it hydrolyzed 4Np-Xyl less efficiently. Activity on 2Np-Cel was observed also in other GH10 enzymes; however, GH10-XA showed marked specificity toward xylan substrates [71].

Supplementary material related to this article found, in the online version, at <http://dx.doi.org/10.1016/j.enzmictec.2015.06.014>

The steady-state kinetic constants of the enzyme are reported in Table 3: the highest specificity constant was observed on MGX followed by 2Np-Cel and beechwood xylan while the enzyme was only barely active on 4Np-Xyl. However, GH10-XA, when incubated in the presence of high concentrations of 4Np-Xyl (30 mM), yielded transglycosylation products observable by TLC (Additional Fig. 5A). Among GH10 enzymes, transglycosylation was only shown for XynB from *Paenibacillus* sp by using xylotriose and xylotetraose substrates [72]. Moreover, it is worth mentioning that, for the first time for GH10, we observed transglycosylation products also in the presence of 2Np-Cel and 4Np-Glc (Additional Fig. 5B). Instead, cel-

Table 2
Comparison of characterized thermophilic GH67.

Microorganism	Mw (kDa)	Optimum temperature (°C)	Optimum pH	Specific activity (U/mg)	Substrate	Reference	Thermostability	Identity vs GH67-GA (%)	Identity vs GH67-GC (%)
<i>Caldicellulosiruptor saccharolyticus</i> (GH67-GC)	109	65	6.5	2.6	Aldouronic acids	this study	50% (65 °C/16 h)–50% (75 °C/2 h)	49	–
<i>Alicyclobacillus acidocaldarius</i> (GH67-GA)	153	50	6.0	4.6	Aldouronic acids	this study	50% (50 °C/30 m)	–	49
<i>Caldicellulosiruptor lactoaceticus</i>	163	75	6.5	1.3	Aldobiuronic acid	[59]	not reported	51	77
<i>Caldicellulosiruptor bescii</i>	not reported	70–75	5.5–6.0	not reported	Not reported	[47]	50% (65 °C/1 m)	50	77
<i>Caldanaerobius polysaccharolyticus</i>	158	60	5.5	154	Aldouronic acids	[64]	not reported	49	59
<i>Uncultured bacterium</i>	not reported	45	7.0–8.0	11.7	Aldouronic acids	[55]	75% (40 °C/120 m)–70% (50 °C/120 m)–0% (60 °C/60 m)	59	58
<i>Geobacillus stereothermophilus</i> T-6	150	65	6.0	42.0	Aldouronic acids	[65]	100% (65 °C/20 m)–100% (70 °C/20 m); 30% (75 °C/20 m)	55	57
<i>Geobacillus stereothermophilus</i> 236	161	40	6.5	14.2	Aldouronic acids	[66]	not reported	54	50
<i>Paenibacillus</i> sp. Strain JDR-2	not reported	40	5.5	5.5	Aldouronic acids	[67]	not reported	54	51
<i>Aureobasidium pullulans</i>	157	65	5.0	14.1 60 135 126 2.3	Aldobiuronic acid Aldotriuronic acid Aldotetrauronic acid Aldopentaauronic acid Aldotriuronic acid	[68]	50% (55 °C/23 m); 50% (65 °C/6 m)	40	37
<i>Schizophyllum commune</i>	330–400	40	5.8	6.2	Aldopentaauronic acid	[32]	65% (45 °C/24 h)	native	native
<i>Pichia stipitis</i>	not reported	60	4.4	1.390 0.06 0.06 0.04 0.08	Aldopentaauronic acid Aldo-bi/triuronic acid Aldobiuronic acid Aldotriuronic acid Aldotetrauronic acid Aldopentaauronic acid	[33]	stable (40 °C/3 h)–50% (60 °C/30 m)	native	native
<i>Phlebia radiata</i>	110	60	3.5			[69]	not reported	native	native
<i>Cellvibrio japonicus</i>	150	37	6.3			[70]	rapidly inactivated at 55 °C	native	native

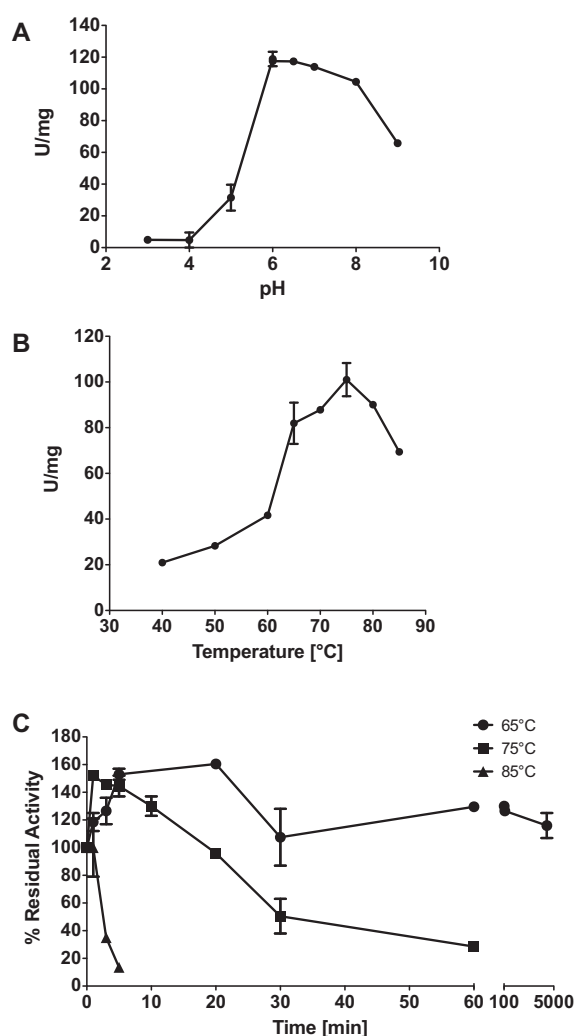


Fig. 1. Characterization of GH10-XA. (A) Specific activity vs pH; the assays were performed in citrate/phosphate and sodium phosphate buffers in the ranges 3.0–6.0 and 6.0–9.0, respectively. (B) Specific activity vs temperature. (C) Thermal stability at 65 °C, 75 °C, and 85 °C (circles, squares, and triangles, respectively).

lobiose and cellotriose were not substrates (Additional Table 4); therefore, GH10-XA performed transglycosylations only from arylglycosides.

Supplementary material related to this article found, in the online version, at <http://dx.doi.org/10.1016/j.enzmictec.2015.06.014>

The mode of action of GH10-XA was tested by performing a time analysis on 2Np-Cel and revealed that the hydrolysis occurred between 2-nitrophenol and cellobiose with the accumulation of the latter after 16 h (Fig. 3A). In addition, GH10-XA degraded XOs primarily into xylobiose and, less efficiently, xylose after 16 h of incubation (Fig. 3B), indicating that GH10-XA removes disaccharidic units from the non-reducing end of the substrate as shown for other GH10 enzymes [73,74].

The hydrolytic activity of GH67-GA and GH67-GC was tested on the substrates listed in Additional Table 4. Both α -glucuronidases were not active on MGX and beechwood xylan but hydrolyzed aldouronic acids. These compounds are a mixture of aldo-tri-, tetra-, and penturonic acids (2:2:1 as relative ratio) showing a 4-O-methyl- α -glucuronic acid bound at the C2 of the xylose residue at the non-reducing end of the xylooligosaccharide (Additional Fig. 6). Reaction mixture containing GH67-GC and aldouronic acids was analysed by HPAEC-PAD (Fig. 4). The analysis of the reaction

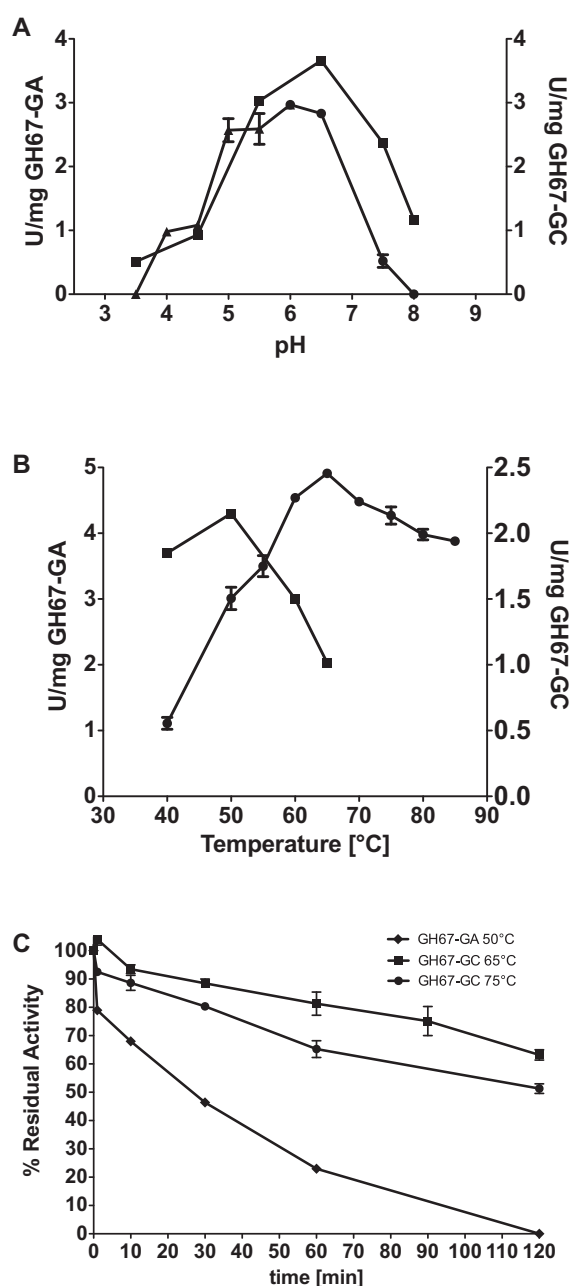


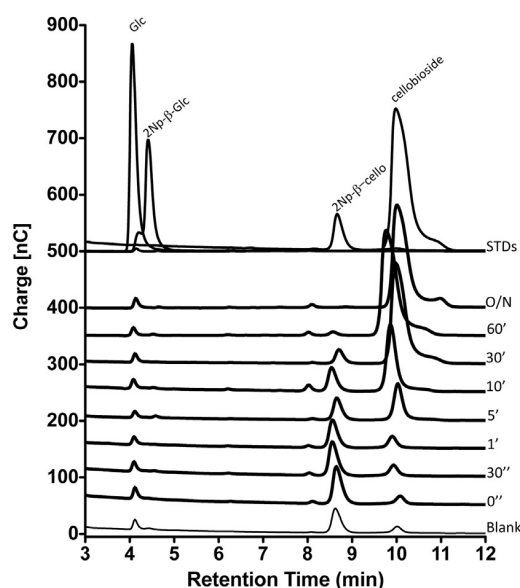
Fig. 2. Characterization of GH67-GA and GH67-GC. (A) Specific activity vs pH, the assays were performed in sodium acetate and in sodium phosphate buffers in the ranges 3.5–5.5 and 5.5–8.0, respectively; squares GH67-GA, triangles GH67-GC. (B) Specific activity vs temperature: square GH67-GA, circles GH67-GC. (C) Thermal stability of GH67-GA at 50 °C and GH67-GC at 65 °C and 75 °C.

products of GH67-GC and of another GH67 α -glucuronidase from *Geobacillus stearothermophilus* (AGUSB) used as positive control, allowed the unequivocal identification of xylose (Xyl1), xylobiose (Xyl2), xylotriose (Xyl3), xylotetraose (Xyl4) oligosaccharides, and of an additional peak eluting at about 15 min of the sodium acetate gradient (Fig. 4). This compound was ascribed to 4-O-methyl- α -glucuronic acid as reported by others [32].

Supplementary material related to this article found, in the online version, at <http://dx.doi.org/10.1016/j.enzmictec.2015.06.014>

The steady-state kinetic constants reported in Table 3 show that the two enzymes have similar k_{cat}/K_M values for aldouronic

A



B

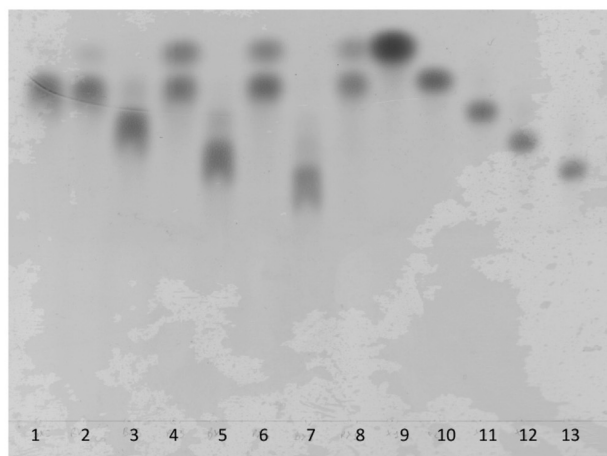


Fig. 3. Analysis of the mode of action of GH10-XA. Time course of the reactions on 2Np-β-cellobioside analysed by HPAEC-PAD (A). TLC analysis of GH10-XA assay on XOIs (B). (1) Blank on xylobiose (Xyl2); (2) reaction on Xyl2; (3) blank on xylotriase (Xyl3); (4) reaction on Xyl3; (5) blank on xylo-tetraose (Xyl4); (6) reaction on Xyl4; (7) blank on xylopentaose (Xyl5); (8) reaction on Xyl5; (9–13) standards: Xyl1, Xyl2, Xyl3, Xyl4, and Xyl5 respectively.

Table 3
Steady-state kinetic constants of GH10 and GH67 enzymes.

GH10-XA			
	k_{cat} (s^{-1})	K_M ($mg\ mL^{-1}$)	k_{cat}/K_M ($s^{-1}\ mg^{-1}\ mL$)
MGX	110.8 ± 5.0	0.79 ± 0.12	140.3
2Np-Cel	56.5 ± 2.1	0.51 ± 0.08	110.8
Beechwood xylan	98.2 ± 4.1	1.81 ± 0.16	54.3
4Np-Xyl	4.4 ± 0.3	1.62 ± 0.30	2.7
GH67-GA			
Aldouronic acids	11.7 ± 0.7	0.23 ± 0.05	51
GH67-GC			
Aldouronic acids	4.8 ± 0.3	0.16 ± 0.03	30
4Np-GlcUA	0.16 ± 0.03	2.93 ± 0.86	0.05

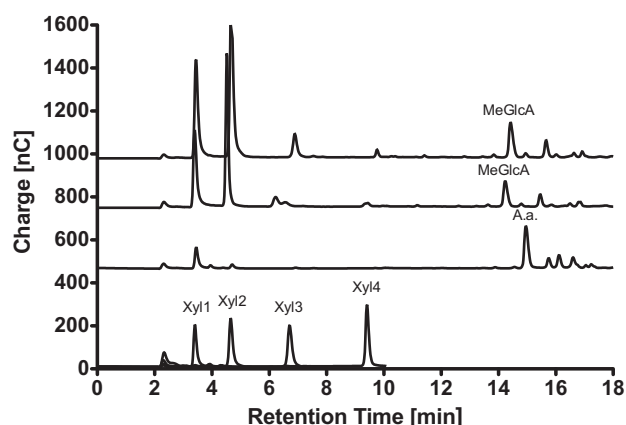


Fig. 4. Hydrolytic activity of GH67-GC on the aldouronic acids substrate. HPAEC-PAD analysis; A.a.: aldouronic acids; MeGlcA: 4-O-methyl glucuronic acid. From bottom to top: standards; blank; reaction with AGUSB; reaction with GH67-GC.

Table 4
Combined effect of GH10, GH67, and GH3 enzymes on glucurono-xylans conversion.

4-O-Methyl-glucurono-xylan		
Enzymes	Products	
	(mM)	(%)
GH10-XA	8.6 ± 0.4	60
GH67-GC	ND	–
GH10-XA + GH67-GC	9.4 ± 0.1	66
GH3-XT	2.1 ± 0.1	15
GH10-XA + GH67-GC + GH3-XT	14.3 ± 0.9	100
Beechwood xylan		
GH10-XA	24.8 ± 0.6	73
GH67-GC	ND	–
GH10-XA + GH67-GC	26.7 ± 0.3	79
GH3-XT	1.1 ± 0.2	3
GH10-XA + GH67-GC + GH3-XT	33.9 ± 0.2	100

acids while it is worth mentioning that GH67-GC was also able to hydrolyze 4-nitrophenyl- α -glucuronide (4Np-GlcUA), albeit less efficiently. This is, at the best of our knowledge, the only α -glucuronidase studied so far that is able to recognize as substrate both glucuronoxylan oligomers and 4Np-GlcUA. In fact, most of the GH67 α -glucuronidases are only active on 4-O-MeGlcA linked to the non-reducing end of xylo-oligomers [29,30,59] and only few studies demonstrated their high activity against longer polymeric substrates [32,33]. By contrast, the Agu4A α -glucuronidase from *Thermotoga maritima* from family GH4, which efficiently hydrolysed 4Np-GlcUA, was unable to recognize glucuronoxylans as substrates [28,75].

Synergy of GH10 and GH67 enzymes on glucuronoxylans. Both 4-O-methyl-glucuronoxylan and beechwood xylan contain 4-O-methyl- α -glucuronide residues bound to the xylan backbone. The substrate specificity of the GH10 and GH67 enzymes presented above prompted us to analyse the effect of the xylanase and of the α -glucuronidases on these substrates.

Preliminary analyses on GH10-XA and of GH67-GA showed that the presence of the latter slightly increased the activity of the xylanase, leading to an increment of the reducing ends of 0.2 μ moles on MGX after 16 h; however, because of the low stability of GH67-GA at these operational conditions, this enzyme was not tested any further. The combined action of GH10-XA and of GH67-GC on MGX and beechwood xylan showed an increment of the reducing ends of 1 and 2 mM, respectively, after only 2 h of incubation (Table 4). The analysis of the reaction products by HPAEC-PAD confirmed the assays and clearly revealed the release of

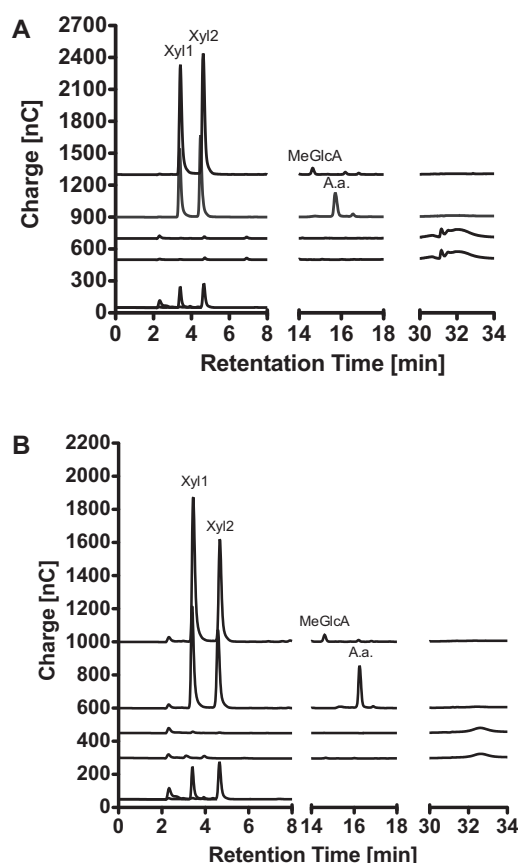


Fig. 5. Analysis of the hydrolytic activity of GH10-XA and GH67-GC on 4-O-methyl glucurono-xylan. The reaction mixtures on MGX (A) and beechwood xylan (B) were analyzed by HPAEC-PAD. From bottom to top: standards; blank; reaction with GH67-GC; reaction with GH10-XA; reaction with GH67-GC and GH10-XA.

xylose, xylobiose, and 4-O-methyl-glucuronic acid: the direct identification, for the first time, of the latter compound confirmed that GH10-XA and GH67-GC cooperate in the hydrolysis of beechwood xylan (Fig. 5). It is worth noting that the hydrolysis of GH10-XA alone accumulated xylobiose, confirming that it was a bad substrate, and a third peak released later than 4-O-methyl-glucuronic acids and possibly corresponding to aldouronic acids.

It is worth mentioning that GH10-XA and GH67-GC acted in synergy extremely well: if compared to the GH10 and GH67 enzymes from *C. lactoaceticus* [59], they produced the same amounts of reducing ends from beechwood xylan in 2 h vs 12 h and at much less enzyme concentrations (7- and 9-fold for GH10 and GH67 enzymes, respectively). On the basis of these encouraging results, and to improve the biotransformation and the synergy of GH10-XA and GH67-GC on these substrates, the high-xylose tolerant β -xylosidase from *Thermotoga thermarum* (GH3-XT) [42] was added to the enzymatic mixture. Interestingly, the reducing ends on both substrates increased substantially in the presence of the three enzymes after only 2 h of incubation (Table 4).

To assess the potential use of the three enzymes in saccharification of plant biomasses we analysed the activities of GH10-XA, GH67-GC and GH3-XT on the liquid fraction of smart cooking pretreated *A. donax*. In fact, it was shown that the hemicelluloses of *A. donax* internodes are constituted by arabino-4-O-methyl-glucuronoxylan [18]. The analysis of the reaction mixtures by HPAEC-PAD revealed the production of xylose as the final product demonstrating that the three enzymes were active on this biomass (Fig. 6A). The absence of 4-O-methyl-glucuronic acid, the other

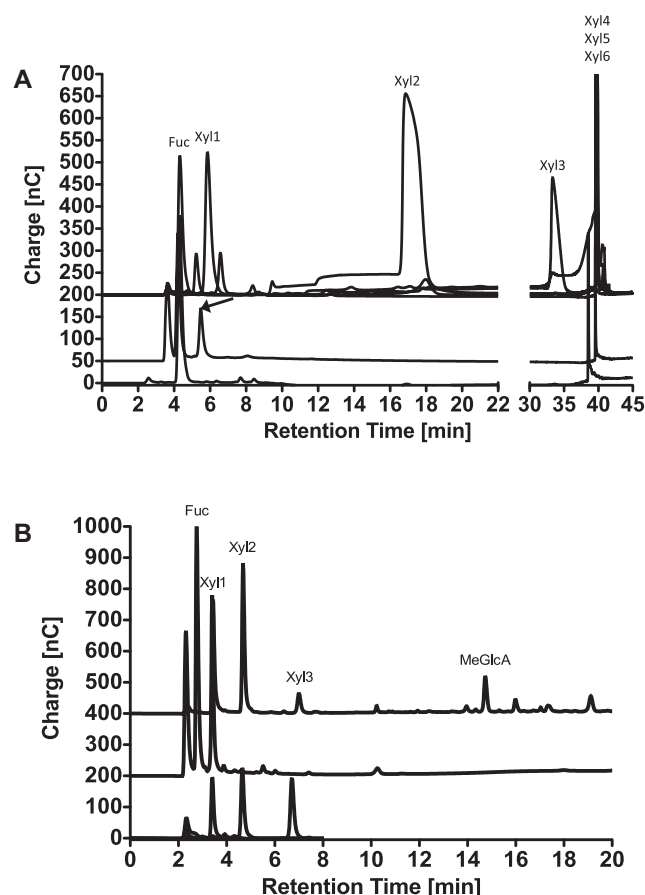


Fig. 6. Analysis of the hydrolytic activity of GH10-XA and GH67-GC on *A. donax* biomass. The reaction mixtures containing GH10-XA, GH67-GC and GH3-XT were analyzed by HPAEC-PAD. (A) From bottom to top: blank; reaction with GH67-GC, GH10-XA and GH3-XT; standards. Fuc: fucose. The arrow indicates the xylose produced while the additional peak in the same chromatogram eluting between 2 and 4 min come from the GH3-XT preparation. (B) The same reaction analyzed with a specific gradient for MeGlcA. From bottom to top: standards; reaction with GH67-GC, GH10-XA and GH3-XT; reaction with AGUSB on aldouronic acids.

expected final product (Fig. 6B) could be due to the low amount of this substituent in the pretreated biomass. These data indicated that GH10-XA, GH67-GC and GH3-XT could be useful components of an enzymatic cocktail for the saccharification of hemicellulosic biomasses for bioethanol production.

4. Conclusions

Carbohydrate active enzymes from extremophilic organisms are considered promising candidates for the preparation of novel enzymatic cocktails for the conversion of lignocellulosic biomasses into fermentable sugars. Despite the relatively easy access to their genes, the characterization of the enzymes for this application is still lagging behind. We showed here the identification through bioinformatic analysis and detailed characterization of GH10 and GH67 hemicellulases from the thermophilic bacteria *A. acidocaldarius* and *C. saccharolyticus* active in synergy on methyl-glucuronoxylans. In particular, GH10-XA and GH67-GC turned out to be substantially more stable to heat than known homologs with identities >50%. In addition, GH10-XA revealed the uncommon ability to catalyze transglycosylation reactions from 4Np-Xyl, while GH67-GC showed remarkable thermal stability and was the first α -glucuronidase known that is able to hydrolyze both aldouronic acids and 4Np-GlcUA. We demonstrated here that

these enzymes, also in the presence of the GH3-XT β -xylosidase, hydrolysed, in synergy and very efficiently in terms of reaction time and enzyme concentrations, glucurono-xylans such as 4-O-methyl-glucuronoxylan, beechwood xylan, and, more importantly, the substrates in the liquid fraction of smart cooking pre-treated biomass from an energy crop. The activity of this enzymatic cocktail on a pretreated biomass coming from a biorefinery production plant is of particular importance as it demonstrates that the use of enzymes from thermophiles is very close to practical application and makes them promising candidates for the treatment of plant biomasses.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

AS carried out the reaction product analysis, participated in the sequence alignment and drafted the manuscript; RI carried out the molecular biology studies, the enzyme characterization, the bio-transformation reactions, and drafted the manuscript; GM carried out the molecular biology studies and the enzyme characterization; RG carried out the molecular biology studies and drafted the manuscript; MR participated in the design of the study; BCP and MM conceived of the study and participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by grant from the Ministero dell'Università e della Ricerca Scientifica -Industrial Research Project "Integrated agro-industrial chains with high energy efficiency for the development of eco-compatible processes of energy and biochemicals production from renewable sources and for the land valorization (Enerbio-Chem)" PON01.01966, funded in the frame of Operative National Programme Research and Competitiveness 2007–2013 D. D. Prot. n. 01/Ric. 18.1.2010. The funding source did not had any role in the study design. We thank Pierpaolo Falcicchio and John van der Oost for the generous gift of the genomic DNA of *C. saccharolyticus* DSM 8903.

References

- [1] D.K.Y. Poon, P. Webster, S.G. Withers, L.P. McIntosh, Characterizing the pH-dependent stability and catalytic mechanism of the family 11 xylanase from the alkalophilic *Bacillus agaradhaerens*, *Carbohydr. Res.* 338 (2003) 415–421.
- [2] A. Ausili, B. Di Lauro, B. Cobucci-Ponzano, E. Bertoli, A. Scire, M. Rossi, et al., Two-dimensional IR correlation spectroscopy of mutants of the beta-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus* identifies the mechanism of quaternary structure stabilization and unravels the sequence of thermal unfolding events, *Biochem. J.* 384 (2004) 69–78.
- [3] D. Shallom, G. Golan, G. Shoham, Y. Shoham, Effect of dimer dissociation on activity and thermostability of the alpha-glucuronidase from *Geobacillus stearothermophilus*: dissecting the different oligomeric forms of family 67 glycoside hydrolases, *J. Bacteriol.* 186 (2004) 6928–6937.
- [4] A. Ausili, B. Cobucci-Ponzano, B. Di Lauro, R. D'Avino, G. Perugini, E. Bertoli, et al., A comparative infrared spectroscopic study of glycoside hydrolases from extremophilic archaea revealed different molecular mechanisms of adaptation to high temperatures, *Proteins* 67 (2007) 991–1001.
- [5] L.D. Unsworth, J.V.D. Oost, S. Koutsopoulos, Hyperthermophilic enzymes; stability, activity and implementation strategies for high temperature applications, *FEBS J.* 274 (2007) 4044–4056.
- [6] S. Atsumi, T. Hanai, J.C. Liao, Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels, *Nature* 451 (2008) 86–89.
- [7] M.E. Himmel, E.A. Bayer, Lignocellulose conversion to biofuels: current challenges, global perspectives, *Curr. Opin. Biotechnol.* 20 (2009) 316–317.
- [8] G. Jager, J. Buchs, Biocatalytic conversion of lignocellulose to platform chemicals, *Biotechnol. J.* 7 (2012) 1122–1136.
- [9] A. Bhalla, N. Bansal, S. Kumar, K.M. Bischoff, R.K. Sani, Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes, *Bioresour. Technol.* 128 (2013) 751–759.
- [10] R. Chen, Y.Z. Wang, Q. Liao, X. Zhu, T.F. Xu, Hydrolysates of lignocellulosic materials for biohydrogen production, *BMB Rep.* 46 (2013) 244–251.
- [11] A. Sunna, M. Moracci, M. Rossi, G. Antranikian, Glycosyl hydrolases from hyperthermophiles, *Extremophiles* 1 (1997) 2–13.
- [12] S.E. Blumer-Schuetz, I. Kataeva, J. Westpheling, M.W. Adams, R.M. Kelly, Extremely thermophilic microorganisms for biomass conversion: status and prospects, *Curr. Opin. Biotechnol.* 19 (2008) 210–217.
- [13] B. Cobucci-Ponzano, V. Aurilia, G. Riccio, B. Henrissat, P.M. Coutinho, A. Strazzulli, A new archaeal beta-glycosidase from *Sulfolobus solfataricus*: seeding a novel retaining beta-glycan-specific glycoside hydrolase family along with the human non-lysosomal glucosylceramidase GBA2, *J. Biol. Chem.* 285 (2010) 20691–20703.
- [14] T. Chang, S. Yao, Thermophilic, lignocellulolytic bacteria for ethanol production: current state and perspectives, *Appl. Microbiol. Biotechnol.* 92 (2011) 13–27.
- [15] D. Barnard, A. Casanueva, M. Tuffin, D. Cowan, Extremophiles in biofuel synthesis, *Environ. Technol.* 31 (2010) 871–888.
- [16] M.E. Himmel, S.Y. Ding, D.K. Johnson, W.S. Adney, M.R. Nimlos, J.W. Brady, et al., Biomass recalcitrance: engineering plants and enzymes for biofuels production, *Science* 315 (2007) 804–807.
- [17] D.B. Jordan, M.J. Bowman, J.D. Braker, B.S. Dien, R.E. Hector, C.C. Lee, et al., Plant cell walls to ethanol, *Biochem. J.* 442 (2012) 241–252.
- [18] C.P. Neto, A. Seca, A.M. Nunes, M.A. Coimbra, F. Domingues, D. Evtuguin, et al., Variations in chemical composition and structure of macromolecular components in different morphological regions and maturity stages of *Arundo donax*, *Ind. Crop Prod.* 6 (1997) 51–58.
- [19] R. Saikia, R.S. Chutia, R. Katak, K.K. Pant, Perennial grass (*Arundo donax* L.) as a feedstock for thermo-chemical conversion to energy and materials, *Bioresour. Technol.* 188 (2015) 265–272.
- [20] A. Teleman, T. Hausalo, M. Tenkanen, T. Vuorinen, Identification of the acidic degradation products of hexenuronic acid and characterisation of hexenuronic acid-substituted xylooligosaccharides by NMR spectroscopy, *Carbohydr. Res.* 280 (1996) 197–208.
- [21] M. Moracci, M. Ciaramella, R. Nucci, L.H. Pearl, I. Sanderson, A. Trincone, et al., Thermostable beta-glycosidase from *Sulfolobus solfataricus*, *Biocatalysis* 11 (1994) 89–103.
- [22] M. Moracci, B. Cobucci-Ponzano, A. Trincone, S. Fusco, M. De Rosa, J. van der Oost, et al., Identification and molecular characterization of the first alpha-xylosidase from an Archaeon, *J. Biol. Chem.* 275 (2000) 22082–22089.
- [23] A. Trincone, B. Cobucci-Ponzano, B. Di Lauro, M. Rossi, Y. Mitsuiishi, M. Moracci, Enzymatic synthesis and hydrolysis of xylogluco-oligosaccharides using the first archaeal alpha-xylosidase from *Sulfolobus solfataricus*, *Extremophiles* 5 (2001) 277–282.
- [24] B. Cobucci-Ponzano, F. Conte, D. Benelli, P. Londei, A. Flagiello, M. Monti, et al., The gene of an archaeal alpha-L-fucosidase is expressed by translational frameshifting, *Nucleic Acids Res.* 34 (2006) 4258–4268.
- [25] J.P. Peacock, J.K. Cole, S.K. Murugapiran, J.A. Dodsworth, J.C. Fisher, D.P. Moser, et al., Pyrosequencing reveals high-temperature cellulolytic microbial consortia in great boiling spring after In Situ lignocellulose enrichment, *PLoS One* 8 (2013).
- [26] S.E. Blumer-Schuetz, S.D. Brown, K.B. Sander, E.A. Bayer, I. Kataeva, J.V. Zurawski, et al., Thermophilic lignocellulose deconstruction, *FEMS Microbiol. Rev.* 38 (2014) 393–448.
- [27] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, *Nucleic Acids Res.* 42 (2014) D490–5.
- [28] C. Suresh, M. Kitaoka, K. Hayashi, A thermostable non-xylanolytic alpha-glucuronidase of *Thermotoga maritima* MSB8, *Biosci. Biotechnol. Biochem.* 67 (2003) 2359–2364.
- [29] D. Nurizzo, T. Nagy, H.J. Gilbert, G.J. Davies, The structural basis for catalysis and specificity of the *Pseudomonas cellulosa* alpha-glucuronidase, GlcA67A, *Structure* 10 (2002) 547–556.
- [30] G. Golan, D. Shallom, A. Teplitsky, G. Zaide, S. Shulami, T. Baasov, et al., Crystal structures of *Geobacillus stearothermophilus* alpha-glucuronidase complexed with its substrate and products - Mechanistic implications, *J. Biol. Chem.* 279 (2004) 3014–3024.
- [31] K.M. Khandke, P.J. Vithayathil, S.K. Murthy, Purification and characterization of an alpha-D-glucuronidase from a thermophilic fungus, *Thermoascus aurantiacus*, *Arch. Biochem. Biophys.* 274 (1989) 511–517.
- [32] M. Tenkanen, M. Siika-aho, An alpha-glucuronidase of *Schizophyllum commune* acting on polymeric xylan, *J. Biotechnol.* 78 (2000) 149–161.
- [33] O. Ryabova, M. Vrsanska, S. Kaneko, W.H. van Zyl, P. Biely, A novel family of hemicellulolytic alpha-glucuronidase, *FEBS Lett.* 583 (2009) 1457–1462.
- [34] T. Collins, C. Gerday, G. Feller, Xylanases, xylanase families and extremophilic xylanases, *FEMS Microbiol. Rev.* 29 (2005) 3–23.
- [35] K. Kolenova, M. Vrsanska, P. Biely, Mode of action of endo-beta-1,4-xylanases of families 10 and 11 on acidic xylooligosaccharides, *J. Biotechnol.* 121 (2006) 338–345.
- [36] K. Eckert, A. Vigouroux, L. Lo Leggio, S. Morera, Crystal structures of A. *acidocaldarius* endoglucanase Cel9A in complex with cello-oligosaccharides: strong -1 and -2 subsites mimic cellobiohydrolase activity, *J. Mol. Biol.* 394 (2009) 61–70.

- [37] B. Di Lauro, M. Rossi, M. Moracci, Characterization of a beta-glycosidase from the thermoacidophilic bacterium *Alicyclobacillus acidocaldarius*, *Extremophiles* 10 (2006) 301–310.
- [38] A.L. Vanfossen, D.L. Lewis, J.D. Nichols, R.M. Kelly, Polysaccharide degradation and synthesis by extremely thermophilic anaerobes, *Ann. N. Y. Acad. Sci.* 1125 (2008) 322–337.
- [39] I. Ozdemir, S.E. Blumer-Schuette, R.M. Kelly, S-layer homology domain proteins CscA.0678 and CscA.2722 are implicated in plant polysaccharide deconstruction by the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*, *Appl. Environ. Microbiol.* 78 (2012) 768–777.
- [40] V.A. Svetlitchnyi, O. Kensch, D.A. Falkenhan, S.G. Korseska, N. Lippert, M. Prinz, et al., Single-step ethanol production from lignocellulose using novel extremely thermophilic bacteria, *Biotechnol. Biofuels* 6 (2013) 31.
- [41] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72 (1976) 248–254.
- [42] H. Shi, X. Li, H.X. Gu, Y. Zhang, Y.J. Huang, L.L. Wang, et al., Biochemical properties of a novel thermostable and highly xylose-tolerant beta-xylosidase/alpha-arabinosidase from *Thermotoga thermarum*, *Biotechnol. Biofuels* 6 (2013).
- [43] N. Nelson, A photometric adaptation of the somogyi method for the determination of glucose, *J. Biol. Chem.* 153 (1944) 375–380.
- [44] M. Somogyi, Notes on sugar determination, *J. Biol. Chem.* 195 (1952) 19–23.
- [45] B. Cobucci-Ponzano, M. Mazzone, M. Rossi, M. Moracci, Probing the catalytically essential residues of the alpha-L-fucosidase from the hyperthermophilic Archaeon *Sulfolobus solfataricus*, *Biochemistry-US* 44 (2005) 6331–6342.
- [46] Y. Bai, J. Wang, Z. Zhang, P. Yang, P. Shi, H. Luo, et al., A new xylanase from thermoacidophilic *Alicyclobacillus* sp. A4 with broad-range pH activity and pH stability, *J. Indus. Microbiol. Biotechnol.* 37 (2010) 187–194.
- [47] X.Y. Su, Y.J. Han, D. Dodd, Y.H. Moon, S. Yoshida, R.I. Mackie, et al., Reconstitution of a thermostable Xylan-degrading enzyme mixture from the bacterium *Caldicellulosiruptor bescii*, *Appl. Environ. Microb.* 79 (2013) 1481–1490.
- [48] C.C. Lee, R.E. Kibblewhite, K. Wagschal, R.P. Li, G.H. Robertson, W.J. Orts, Isolation and characterization of a novel GH67 alpha-glucuronidase from a mixed culture, *J. Indus. Microbiol. Biotechnol.* 39 (2012) 1245–1251.
- [49] P. Ruile, C. Winterhalter, W. Liebl, Isolation and analysis of a gene encoding alpha-glucuronidase, an enzyme with a novel primary structure involved in the breakdown of xylan, *Mol. Microbiol.* 23 (1997) 267–279.
- [50] B. Cobucci-Ponzano, A. Trincone, A. Giordano, M. Rossi, M. Moracci, Identification of an archaeal alpha-L-fucosidase encoded by an interrupted gene. Production of a functional enzyme by mutations mimicking programmed -1 frameshifting (vol 278, pg 14,622, 2003), *J. Biol. Chem* 278 (2003) 47350.
- [51] A. Sunna, M.D. Gibbs, P.L. Bergquist, A novel thermostable multidomain 1,4-beta-xylanase from 'Caldibacillus cellulovorans' and effect of its xylan-binding domain on enzyme activity, *Microbiology* 146 (Pt (11)) (2000) 2947–2955.
- [52] A. Khasin, I. Alchanati, Y. Shoham, Purification and characterization of a thermostable xylanase from *Bacillus stearothermophilus* T-6, *Appl. Environ. Microbiol.* 59 (1993) 1725–1730.
- [53] H.D. Simpson, U.R. Haufler, R.M. Daniel, An extremely thermostable xylanase from the thermophilic eubacterium *Thermotoga*, *Biochem J.* 277 (Pt (2)) (1991) 413–417.
- [54] Z.Q. Jiang, W. Deng, X.T. Li, Z.L. Ai, L.T. Li, I. Kusakabe, Characterization of a novel, ultra-large xylanolytic complex (xylanosome) from *Streptomyces olivaceoviridis* E-86, *Enzyme Microb. Tech.* 36 (2005) 923–929.
- [55] F. Zheng, J.X. Huang, Y.H. Yin, S.J. Ding, A novel neutral xylanase with high SDS resistance from *Volvariella volvacea*: characterization and its synergistic hydrolysis of wheat bran with acetyl xylan esterase, *J. Indus. Microbiol. Biotechnol.* 40 (2013) 1083–1093.
- [56] J. Zheng, N. Guo, L.S. Wu, J. Tian, H.B. Zhou, Characterization and constitutive expression of a novel endo-1,4-beta-D-xylanohydrolase from *Aspergillus niger* in *Pichia pastoris*, *Biotechnol. Lett.* 35 (2013) 1433–1440.
- [57] H.Y. Luo, J. Yang, J. Li, P.J. Shi, H.Q. Huang, Y.G. Bai, et al., Molecular cloning and characterization of the novel acidic xylanase XYLD from *Bispora* sp MEY-1 that is homologous to family 30 glycosyl hydrolases, *Appl. Microbiol. Biotechnol.* 86 (2010) 1829–1839.
- [58] K.S. Hung, S.M. Liu, W.S. Tzou, F.P. Lin, C.L. Pan, T.Y. Fang, et al., Characterization of a novel GH10 thermostable, halophilic xylanase from the marine bacterium *Thermoanaerobacterium saccharolyticum* NT0U1, *Process Biochem.* 46 (2011) 1257–1263.
- [59] X.J. Jia, S.F. Mi, J.Z. Wang, W.B. Qiao, X.W. Peng, Y.J. Han, Insight into glycoside hydrolases for debranched xylan degradation from extremely thermophilic bacterium *Caldicellulosiruptor lactoaceticus*, *PLoS One* 9 (2014).
- [60] M.P. Sapre, H. Jha, M.B. Patil, Purification and characterization of a thermostable-cellulase free xylanase from *Syncephalastrum racemosum* Cohn, *J. Gen. Appl. Microbiol.* 51 (2005) 327–334.
- [61] Z. Taibi, B. Saoudi, M. Boudelaa, H. Trigui, H. Belghith, A. Gargouri, et al., Purification and biochemical characterization of a highly thermostable xylanase from *Actinomadura* sp strain Cpt20 isolated from poultry compost, *Appl. Biochem. Biotechnol.* 166 (2012) 663–679.
- [62] G.C. Pradeep, Y.H. Choi, Y.S. Choi, C.N. Seong, S.S. Cho, H.J. Lee, et al., A novel thermostable cellulase free xylanase stable in broad range of pH from *Streptomyces* sp. CS428, *Process Biochem.* 48 (2013) 1188–1196.
- [63] A. Gessesse, Purification and properties of two thermostable alkaline xylanases from an alkaliphilic *Bacillus* sp, *Appl. Environ. Microb.* 64 (1998) 3533–3535.
- [64] Y. Han, V. Agarwal, D. Dodd, J. Kim, B. Bae, R.I. Mackie, et al., Biochemical and structural insights into xylan utilization by the thermophilic bacterium *Caldanaerobius polysaccharolyticus*, *J. Biol. Chem.* 287 (2012) 34946–34960.
- [65] G. Zaide, D. Shallom, S. Shulami, G. Zolotnitsky, G. Golan, T. Baasov, et al., Biochemical characterization and identification of catalytic residues in alpha-glucuronidase from *Bacillus stearothermophilus* T-6, *Eur. J. Biochem.* 268 (2001) 3006–3016.
- [66] I.D. Choi, H.Y. Kim, Y.J. Choi, Gene cloning and characterization of alpha-glucuronidase of *Bacillus stearothermophilus* No. 236, *Biosci. Biotechnol. Biochem.* 64 (2000) 2530–2537.
- [67] G. Nong, J.D. Rice, V. Chow, J.F. Preston, Aldouronate utilization in *Paenibacillus* sp strain JDR-2: physiological and enzymatic evidence for coupling of extracellular depolymerization and intracellular metabolism, *Appl. Environ. Microb.* 75 (2009) 4410–4418.
- [68] B.J.M. de Wet, W.H. van Zyl, B.A. Prior, Characterization of the *Aureobasidium pullulans* alpha-glucuronidase expressed in *Saccharomyces cerevisiae*, *Enzyme Microb. Technol.* 38 (2006) 649–656.
- [69] M. Mierzwa, J. Tokarzewska-Zadora, T. Deptula, J. Rogalski, J. Szczodrak, Purification and characterization of an extracellular alpha-D-glucuronidase from *Phlebia radiata*, *Prep. Biochem. Biotechnol.* 35 (2005) 243–256.
- [70] T. Nagy, K. Emami, C.M.G.A. Fontes, L.M.A. Ferreira, D.R. Humphry, H.J. Gilbert, The membrane-bound alpha-glucuronidase from *Pseudomonas cellulosa* hydrolyzes 4-O-methyl-D-glucuronoxyloligosaccharides but not 4-O-methyl-D-glucuronoxylan, *J. Bacteriol.* 184 (2002) 4925–4929.
- [71] D.Y. Kim, D.H. Shin, S. Jung, H. Kim, J.S. Lee, H.Y. Cho, et al., Novel Alkali-Tolerant GH10 Endo-beta-1,4-Xylanase with broad substrate specificity from microbacterium *trichothecenolyticum* HY-17, a gut bacterium of the mole cricket *Gryllotalpa orientalis*, *J. Microbiol. Biotechnol.* 24 (2014) 943–953.
- [72] A. Blanco, P. Diaz, J. Martinez, O. Lopez, C. Soler, F.I. Pastor, Cloning of a *Bacillus* sp. BP-23 gene encoding a xylanase with high activity against aryl xylosides, *FEMS Microbiol. Lett.* 137 (1996) 285–290.
- [73] C.M. Fontes, H.J. Gilbert, G.P. Hazlewood, J.H. Clarke, J.A. Prates, V.A. McKie, et al., A novel *Cellvibrio mixtus* family 10 xylanase that is both intracellular and expressed under non-inducing conditions, *Microbiology* 146 (Pt (8)) (2000) 1959–1967.
- [74] K. Usui, T. Suzuki, T. Akisaka, K. Kawai, A cytoplasmic xylanase (XynX) of *Aeromonas caviae* ME-1 is released from the cytoplasm to the periplasm by osmotic downshock, *J. Biosci. Bioeng.* 95 (2003) 488–495.
- [75] C. Suresh, A.A. Rus'd, M. Kitaoka, K. Hayashi, Evidence that the putative alpha-glucosidase of *Thermotoga maritima* MSB8 is a pNP alpha-D-glucuronopyranoside hydrolyzing alpha-glucuronidase, *FEBS Lett.* 517 (2002) 159–162.

~~~~~Conferences~~~~~

2016

Novel thermophilic hemicellulases for the conversion of lignocellulose for second generation biorefineries. Iacono R., Cobucci-Ponzano B., Strazzulli A., Masturzo G., Giglio R., Rossi M., Moracci M. Proteine, March 30st- April 1st 2015, Bologna, Italy.

2015

*Identification and characterization of a novel unclassified de-N-acetylase from *Sulfolobus solfataricus*.* Iacono R., Cobucci- Ponzano B., Strazzulli A., Moracci M. 11th Carbohydrate Bioengineering Meeting, May 10-13th 2015, Helsinki, Finland.

Discovery of novel Carbohydrate Active Enzymes for plant biomass degradation by metagenomics of hyperthermophilic communities. Cobucci-Ponzano B., Strazzulli A., Giglio R., Iacono R., Bitetti F., Schiano di Cola C., Lauro FM., Zhou Y., Xu J., Lombard V., Henrissat B., Cardoso V., Fontes CMGA and Moracci M. 11th Carbohydrate Bioengineering Meeting, May 10-13th 2015, Helsinki, Finland.

2014

Stabilization of the human lysosomal α -glucosidase by allosteric chaperones for the treatment of Pompe disease. Cobucci-Ponzano B., Ferrara MC., Porto C., Iacono R., Meli4M., Colombo G., Parenti G., Moracci M. Meeting IBBR October 6-7th 2014, Naples, Italy.

Exploring and exploiting by metagenomics the biodiversity of hyperthermophilic communities for novel enzymes for plant biomass biodegradation. Strazzulli A., Cobucci-Ponzano B., Giglio R., Iacono R., Zhou Y., Xu J., Lauro F., Henrissat B. and Moracci M. 7th international Congress on Biocatalysis, August 31st- September 4th 2014, Hamburg, Germany.

2013

Identification of new thermophilic enzymatic activities for the biotransformation of plant biomasses. Strazzulli A., Cobucci-Ponzano B., Giglio R., Iacono R., Masturzo G., Zhou Y., Xu J., Lauro F., Henrissat B. and Moracci M.

10th Carbohydrate Bioengineering Meeting, April 21-24th 2013, Prague, Czech Republic.

2012

New extremophilic biocatalysts for the biotransformation of plant biomasses, Moracci M., Cobucci Ponzano B., Morana A., Ionata E., Strazzulli A., Giglio R., Ferrara M.C., Del Monaco G., Maurelli L., Marcolongo L., Iacono R., Masturzo G., La Cara F., Italian Forum on industrial Biotechnology and Bioeconomy, October 23-25th, Milan, Italy.

Identification of new thermophilic enzymatic activities for hemicellulose conversion, Cobucci-Ponzano B., Strazzulli A., Giglio R., Iacono R., Masturzo G., Fagnano M., Moracci M., 56th National Meeting of the Italian Society of Biochemistry and Molecular Biology, September 26-29th, Chieti, Italy.

~~~~~Acknowledgments~~~~~

This work would not have been possible without the support and guidance that I received during my PhD.

I would sincerely thank to my research tutor Dr Marco Moracci, for giving me the opportunity to perform my formation activities in his laboratory, for his valuable guidance and for his consistent encouragement. I am extremely grateful to Dr Beatrice Cobucci Ponzano for her constant supporting and for opening my mind with her precious advices. A special thank to my colleague and friend Dr Andrea Strazzulli for our nice and productive “discussions” in lab and for his constant helpful. I would thank my dear friend Dr Carmen Ferrara for nice moments spent inside and outside lab, for teaching me to be decisive in my aims. Thank to Giuseppe, Rosa, Federica, Luisa, Nicola with whom I shared good times in lab.

A huge thank to Alessandro for his continued support, his patience and to be always present for me. Thanks to my wonderful family for its "cheering from the stadium" to my every goal achieved.

Also thanks to each person who in his own way has contributed to this work.